

# THIS WEEK

## EDITORIALS

**LOST PROPERTY** Discovery on Mars ends long search for Beagle 2 **p.412**

**WORLD VIEW** Managing the terror risk demands complex thought **p.413**

**MINERAL** Earth's nitrogen traced to ice from early Solar System **p.414**



## Kept on a leash

*A vital dependence of genetically modified organisms on an artificial nutrient could be a means of preventing their escape into the environment.*

Of all the concerns and complaints about the use of genetically modified organisms (GMOs), one of the most legitimate has always been the threat of pollution. Life will, after all, find a way. And genetically altered life is no different. From super-weeds made resistant to herbicides, to contamination of food with genetically modified ingredients, it has proved difficult to keep the technology fully contained.

The risks posed by the unintended spread of GMOs are uncertain, but they are infinitely lower than the nightmare scenarios painted by opponents. Yet in many respects that is beside the point. As this journal has pointed out before, public opinion matters here, whether or not scientists think that the basis for that opinion is misguided. Researchers who would create and release GMOs have a duty — ethical and often statutory or regulatory — to minimize the collateral impact on those who do not share their enthusiasm for the technology. Most, of course, already take that responsibility seriously. But many of the promised benefits of genetic modification, from the development of new fuels and medicines to more-durable crops, demand broad deployment outside the laboratory. And outside the laboratory, things can and do go wrong.

If genetic modification cannot be kept in its box, then can it be kept on a leash? Two papers published this week on *Nature's* website offer hope that it can be. In principle, the techniques described in the papers could defuse the major concerns about GMO pollution. But, depending on how the science is applied, they could exaggerate some of the separate worries about the social and economic impact that continue to dog the commercial application of genetic modification, in Europe in particular.

The new approach gives GMOs an Achilles heel. The researchers who have produced the organism have built in vital dependency on an artificial nutrient. If the nutrient is withdrawn, or the organism spreads to where it is no longer available, then the organism cannot survive. Pull too tightly on the leash and it turns into a noose.

The research marks an elegant step forward for the growing field of synthetic biology. In the first paper, Farren Isaacs and his colleagues at Yale University in New Haven, Connecticut, describe how they have produced various GMOs whose growth is restricted by the expression of multiple essential genes that depend on synthetic amino acids (A. J. Rovner *et al.* *Nature* <http://dx.doi.org/10.1038/nature14095>; 2015). In the second, separate study, George Church at Harvard Medical School in Boston, Massachusetts, and his colleagues redesigned essential enzymes in a GMO to make it metabolically dependent on synthetic amino acids (D. J. Mandell *et al.* *Nature* <http://dx.doi.org/10.1038/nature14121>; 2015). The modifications are made throughout the genome to make it harder for the altered sequences to be ejected.

More details of the science are included in a News story on page 423. In both cases, the organisms cannot use alternative nutrients as a

supplement when the synthetic amino acids are unavailable, and they seem to be resistant to mutations that improve their survival in the wild, thanks to the multiple sites of genetic modification. In theory, these features make the control of GMOs more effective than with current techniques. The research in both papers is with bacteria, but there seems no reason why the techniques they describe could not be used to engineer more-complex, multicellular organisms — including crops — in the same way.

***“It is important not to gloss over concerns, however irrational they may seem to some.”***

So what is the downside? Much of the controversy over genetic modification relates to early, clumsy, attempts by big business to commercialize crops, and to gain control over where, when and how they were grown to maximize profit. A crop that needs constant nourishment with a bespoke foodstuff — unavailable elsewhere and with manufacture protected under probable patents — could be presented as a way of tying vulnerable farmers still closer to largely unloved seed companies.

It is early days, but is it too early to consider such a scenario? Yes and no. As the researchers correctly point out, the technology could most immediately be useful for containing GMOs used in research and industrial systems — microbes typically — to make high-value pharmaceuticals, for example. But applications in crops are some way off.

Still, perceptions matter, and it is important not to gloss over concerns, however irrational they may seem to some. Last week, the European Union agreed on compromise legislation that might settle more than a decade of disagreement and allow research and development on GMOs to proceed (see *Nature* **516**, 143; 2014). The new research can address an old fear — contamination — that contributed to the controversy. But it is prudent to consider other anxieties that might sprout in its place. ■

## Down to earth

*A concerted focus on soils will benefit society in untold ways and should be embraced.*

The year 2015 carries many titles. For physics fans, it is the International Year of Light, marking 100 years since Albert Einstein developed his general theory of relativity. For development experts, it is the International Year of Evaluation, a time to promote evidence-based policy-making.

But 2015 also bears earthier significance. The Food and Agriculture Organization of the United Nations (FAO) has decreed

it the International Year of Soils. Britain's leading Earth-sciences organization, the Geological Society in London, has also proclaimed it the Year of Mud. Just add rain. And, as visitors to *Nature's* national and spiritual home can very often attest, what is soil to the world is mud in Britain.

Wet or dry, such dirty pursuits offer an important opportunity to highlight an oft-forgotten corner of the research world. Mud and soil science underlie many of the questions facing our planet, from how to feed billions of people to how to keep them safe from floods, fires and other natural disasters. *Nature* has long argued that researchers must be on the front lines of turning soil discoveries into practical advances. In 2012, one of our journalists travelled to Malawi to analyse the best ways to enrich African soil (see *Nature* **483**, 525–527; 2012). Last week, we assessed the often-overhyped promise of biochar, burnt agricultural material meant to improve the performance of soils that it is mixed into (see *Nature* **517**, 258–260; 2015).

The goal of the Year of Soils is to make others aware as well, and it offers a chance to marshal research initiatives to answer some of the most pressing questions. For instance, the FAO is focusing on updating and revising soil maps around the globe. 'Legacy' maps from Afghanistan to Zimbabwe have been scanned and made available online, providing key historical data for understanding how land use has changed over time. Looking ahead, an international consortium is working on producing a digital global soil map, georeferenced down to 100 metres.

Figuring out which soils lie where is just the first step; after that comes saving the best of what is left. The FAO estimates that an area of soil the size of Costa Rica is lost every year to factors such as erosion, compaction and salinization. One-third of the world's soils have already been degraded, often in the countries that can least afford to lose that resource. When soil becomes sicker, so too do the people who rely on it. Contamination soars and crop yields and human health decline.

Somehow, researchers must also figure out a way for farmers to feed ever-increasing populations on ever-shrinking amounts of arable land. The solutions must take cost into account, because what works in the lab will not work in the field if the soil-treatment or conservation approach is too expensive for the people who need it most.

Against such enormous societal problems, the Geological Society's decision to focus on mud might at first elicit a snigger: mud is a problem to face with wellies, not with a global research agenda. But take a moment to consider how mud influences the world of science.

***"When soil becomes sicker, so too do the people who rely on it."***

For starters, there is the entire record of the history of life. Fossils preserved in mudstone, such as the exquisite Burgess Shale high in the Canadian Rockies, reveal the story of vertebrate evolution. Without small creatures drowning in and being encased by mud, we would have a much harder time unravelling the relationship between organisms past and present.

Then there is the economic importance of mud-based rocks. Petroleum engineers have been exploring shale as a future source of both oil and natural gas. Although controversy rages about how much shale gas might ultimately be available (see *Nature* **516**, 28–30; 2014), extraction rates have soared in the United States, driven by big reserves such as the Marcellus Shale underlying much of Pennsylvania and neighbouring states. In March, industry and academic experts will gather in London to assess the numbers behind a possible ramp-up in shale-resource production in the United Kingdom.

Finally, consider how soil and mud combine to underpin many of the globe's natural disasters. Assessing flood risk requires knowing what soils are where, and how likely they are to turn to mud in times of heavy downpours.

Now is the time to drill down hard on the big questions for soil science. Not all annual focuses are as relevant for society: 2016 has already been designated the International Year of Camelids. ■

## Lost and found

*The discovery of Beagle 2 on Mars should spur the search for other items lost to science.*

**A**n odd idiom in the English language points out that when an object is mislaid, it will turn up "always in the last place you look for it". Space scientists never stopped looking for Beagle 2, the probe that vanished when British researchers tried to land it on Mars in 2003, and last week their persistence paid off. A tiny, shiny smudge in a high-resolution image of the red planet, taken by NASA's Mars Reconnaissance Orbiter (MRO), is very probably the missing spacecraft, the researchers told a press conference in London.

The blurry pictures suggest that the shoestring mission, costing some £50 million (US\$76 million), did better than many have suggested. It is impossible to be sure, but the craft's complicated entry, descent and landing sequence — which involved a parachuted descent followed by a series of bounces cushioned by airbags — seems to have worked as planned. Landing intact on the planet's surface, the probe then probably failed to open all of its solar panels properly, leaving it unable to call home. The exact reason for this will probably remain a mystery. Perhaps the lander hit a rock, or arrived at a faster pace than it should have (data suggest the planet's atmosphere was thinner than expected), or maybe one of the airbags failed to detach and got in the way.

If the British team was unlucky then, it is fortunate now. When the press conference was announced — but not the reason for it — some commentators scoffed at the suggestion that the spacecraft had been found. Measuring just 2 metres across, Beagle 2 was found stranded

in a landing zone some 60 kilometres long and 10 kilometres wide. Although the scientists have examined photos of the Martian surface for years, the find only became a realistic possibility with the 2006 arrival in orbit of the MRO's HiRISE camera, which has a pixel resolution of about 25 centimetres.

The discovery, and the signs that the probe fell only at the final hurdle, marks a victory of sorts for the mission, which was launched by a group of British universities and headed by the media-savvy researcher Colin Pillinger, who died last year. Some space scientists who were critical of the low-budget mission unfairly confused the team's enthusiasm and determination for amateurism — including the European Space Agency, which published a scathing report on the lander's loss in 2004. Bad luck can skewer even the most costly space missions.

At the very least, the happy ending to the decade-long search should encourage others to keep looking for their own missing equipment, artefacts and specimens. For a precise pursuit, science can be surprisingly sloppy. It is easy, one can concede, to get lost in space, but we should be more careful on Earth. NASA, for example, complained in 2012 that dozens of pieces of its Moon rock, brought back by Apollo astronauts and distributed along with other space samples as tokens of goodwill to US states and foreign nations, have been mislaid. Fragments of meteorites have been lost in the post, and a valuable chunk of Moon collected by the *Apollo 11* mission is believed to have been discarded on a rubbish tip in Ireland.

Last month, the University of Texas at Austin said that it was searching for 100 human brains that had been lent to its psychology department by a local hospital in the 1980s, but were now unaccounted for. One thing is certain: they will be in the last place the university thinks to look. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunq](http://go.nature.com/xhunq)





## Effective risk response needs a prepared mindset

**Leaders and risk strategists must collaborate effectively on decisions of global importance, says Erwann Michel-Kerjan.**

20 January 2015

As the global reaction to the terror attacks in Paris shifts from shock and grief to questioning, politicians in France and abroad are seeking ways to minimize the risk of similar atrocities in future.

Post-crisis situations are typically handled badly. Politicians pander to the demand for 'something to be done' (or seen to be done). Short-term solutions are rushed in to address the crisis, until public attention switches to something else. There are already signs of political fire-fighting in the wake of the Paris attacks — British Prime Minister David Cameron has talked of outlawing electronic encryption. True, this might make life more difficult for a small number of would-be terrorists. But the knock-on effect on broader society of restricting this technology would cause more problems than it solves.

### Related stories

- Psychologists seek roots of terror
- Attempts to predict terrorist attacks hit limits
- Science and satire

[Print](#)

The days of quick fixes to risk are over. The key now is to integrate advances in risk management into decision-making, and doing so (or not) has already affected outcomes of several national elections. The competent management of the Hurricane Sandy disaster helped US President Barack Obama's 2012 re-election, whereas the 2004 Spanish elections were lost by the government after it dealt poorly with the Madrid terrorist attacks in March that year. The Paris attacks are high on the agenda at the summit of the World Economic Forum this week in Davos, Switzerland.

A systemic approach is necessary. A simplistic take on the Paris event, for instance, would be that two brothers were radicalized, and killed cartoonists who made fun of the Prophet Muhammad. This might be true, but it is not very helpful.

Advanced risk management quickly connects several trends, often misperceived as independent. In France, these include unsolved immigration issues and reduction in the police workforce — a direct consequence of the current fiscal crisis. Thousands of kilometres away, they include the geopolitical conflicts in the Middle East that give rise to uncontrolled extreme jihadism. Killing journalists allows terrorists to use today's 24/7 media coverage as a social amplifier to spread fear globally. Economic, geopolitical, societal and technological factors reinforced each other.

Beyond terrorism, recent years have seen a series of events caused by converging trends that start in one part of the planet and quickly have far-reaching effects: social unrest, cyberattacks and espionage, climate change, natural disasters, water scarcity, pandemics, economic crises, rising inequality and so on.

It is no surprise, then, that there is now high demand for strategic risk-management expertise in world leaders' offices and corporate boardrooms. But how to foster an effective collaboration? Modern risk management builds on an innovative process with six core stages.

First, do not consider only the risk you know best. Experts in one type of risk — radicalization, for example — tend to overlook the importance of other factors. To avoid this, consult widely.

Second, assess quality and quantity. Computer models can perform advanced quantitative analysis, but emotions and perceptions still influence decision-making. Find a balance and acknowledge uncertainty.

**“There is now high demand for risk-management expertise in world leaders’ offices.”**

Third, do not approach risks in a silo; map their interdependencies. Monitor hazard trends that can reinforce each other.

Fourth, think beyond tomorrow. Short-term thinking is natural for politicians with an eye on the next election or business leaders concerned with next quarter’s results. Risks do not materialize overnight. Look at evolving trends over the past ten years and expected ones for the next ten.

Fifth, rinse and repeat. Risk analysis benefits from re-evaluation, say every year. Bring specialists together regularly to discuss findings openly.

Sixth, communicate risk effectively. For scientists, avoid jargon. Don’t say, “It depends.” Instead, say, “Here are three possible outcomes based on three different assumptions.”

Is that all feasible? My experience in the past decade shows that it is. Since 2005, our Wharton Risk Center team in Philadelphia, Pennsylvania, has collaborated with the World Economic Forum and other partners to produce an influential report ahead of Davos to better foresee such risks.

The Global Risks Report, now in its tenth year, groups 30–50 global risks into five categories: economic, environmental, geopolitical, societal and technological. It compares their likelihood and severity on a ten-year horizon and builds a map of their interdependencies. It combines results of a risk-perception survey of nearly 1,000 leaders in academia, business, government and non-profit circles with technical assessments and several workshops during the year. The findings are written in non-technical terms and accessible to all on the Internet, from global leaders to concerned citizens.

A leading risk from the 2007 report was ‘asset price collapse’ — an unpopular forecast at a time of great exuberance. A year later, we all witnessed the systemic downfall of the financial system.

Advances in risk management might not prevent all crises, but they do help to build a more pragmatic mindset to mitigate risks and strengthen resilience. Interstate conflict tops the 2015 ranking. This is no coincidence.

*Nature* **517**, 413 (22 January 2015) doi:10.1038/517413a

See News pages 419 and 420

## Related stories and links

### From nature.com

- **Psychologists seek roots of terror**  
20 January 2015
- **Attempts to predict terrorist attacks hit limits**  
20 January 2015
- **Science and satire**  
13 January 2015

## Author information

### Affiliations

Erwann Michel-Kerjan is executive director of the Wharton Risk Center at the University of Pennsylvania in Philadelphia, USA, and a member of the World Economic Forum Global Risks Advisory Board.

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## BIOENGINEERING

### Muscle in a dish twitches

Human muscle that contracts has been grown in the lab.

Existing models of human skeletal muscle are two-dimensional and do not mimic the structure or behaviour of natural tissue well. Nenad Bursac at Duke University in Durham, North Carolina, and his colleagues took samples of living human muscle cells and grew them in three dimensions using a scaffold. This coaxed the cells into forming muscle that could spontaneously twitch. When the team stimulated it with electrical pulses similar to nerve signals, the muscle contracted.

The tissue also responded to various drugs, including a steroid-like one, in much the same way as human muscles. The researchers plan to use their tissue to test drugs for muscle disorders.

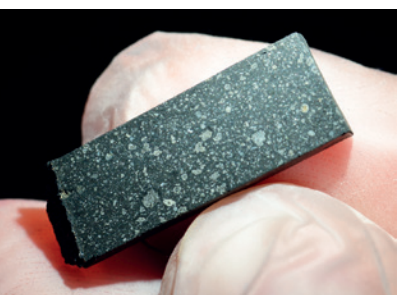
*eLife* <http://doi.org/zfs> (2015)

## COSMOCHEMISTRY

### How nitrogen got to Earth

Earth's nitrogen may have originated in the icy reaches of the primordial Solar System.

A team led by Dennis Harries of the University of Jena in Germany discovered and analysed a chromium nitride mineral inside two meteorites (pictured). The



researchers say that the mineral could have formed from ammonia in ices that swirled around the newborn Sun. Shock waves from distant collisions between fragments in this protoplanetary disk heated up the ammonia, releasing it to react with metals such as chromium. The team also found that the isotopic signature of the nitrogen in the mineral is similar to that of Earth's atmospheric nitrogen.

Much of the early Solar System's nitrogen could have been released in this way, with some ending up in Earth's early atmosphere, the authors suggest.

*Nature Geosci.* <http://dx.doi.org/10.1038/ngeo2339> (2015)

## ANIMAL BEHAVIOUR

### Turtles' magnetic attraction to home

Sea turtles use geomagnetic signatures to return to nesting sites near where they were born.

These animals navigate across oceans using Earth's magnetic field, but it has been unclear how they find the same coastal nesting sites as their mothers. Roger Brothers and Kenneth Lohmann at the University of North Carolina in Chapel Hill studied loggerhead sea turtles (*Caretta caretta*; pictured) in Florida.

They found that the location of the animals'

nests each season was associated with changes in the strength and direction of Earth's magnetic field at each site. In areas where the same magnetic signature spreads out over time, nests were made farther apart. When the signature shrank, the nests were closer together.

Similar mechanisms could be at work in other animals that migrate back to their birthplaces, the authors suggest.

*Curr. Biol.* <http://doi.org/zc7> (2015)



## DEVELOPMENTAL BIOLOGY

### Stem cells for bone growth

Stem cells that give rise to bone and cartilage in mice after birth have been found by two teams.

A group led by Siddhartha Mukherjee and Timothy Wang at Columbia University in New York found that cells at the ends of mouse bones can produce other cells that make bone, cartilage and the spongy tissue in bone marrow. When implanted near a broken bone, these stem cells developed into bone-making cells.

Charles Chan and Michael Longaker at Stanford University in California and

their colleagues identified mouse stem cells with similar capabilities, as well as the molecular signals that maintain such cells and guide their development. When combinations of these factors were added to fat tissue in mice along with collagen protein, bone or cartilage formed a month later.

*Cell* 160, 269–284; 285–298 (2015)

## ECOLOGY

### Small trees save forests

Small trees are often removed from conifer forests in dry areas to reduce the risk of wildfires, but a US study has

J. ROGER BROTHERS

DENNIS HARRIES



revealed that insect outbreaks pose a greater threat to such forests than fire does.

William Baker and Mark Williams at the University of Wyoming in Laramie studied records of fires and insect outbreaks in dry forests in the western United States from 1999 to 2012, and found that insects caused the loss of 5.6 times more forest area than did wildfires. The team also analysed land surveys from the late nineteenth century and found that small trees — of about 40 centimetres or less in diameter — made up almost 62% of total trees, a higher proportion than was thought.

The findings suggest that, historically, forests were resilient because of the smaller trees, and that removing these to control fire risk could be misguided.

*Front. Ecol. Evol.* <http://doi.org/zfj> (2015)

## NEUROSCIENCE

## How baby rodents block pain

Nerve injury in very young animals does not result in pain as it does later in life, probably because of an anti-inflammatory response in the spinal cord.

Maria Fitzgerald of University College London and her colleagues damaged hind-limb nerves in rat and mouse pups. They tested the sensitivity of the paws to painful stimuli, such as mechanical pressure, then recorded the excitability of spinal neurons at different ages and analysed immune profiles. They found that both pain sensitivity and neural excitability developed in the pups at an age equivalent to human adolescence, when anti-inflammatory molecules were replaced by those that promote inflammation.

The team reversed this pain suppression in injured young rats by blocking production of IL-10, an anti-inflammatory cytokine. The results could explain why such 'neuropathic' pain sometimes emerges

mysteriously in human adolescents, the authors say. *J. Neurosci.* 35, 457–466 (2015)

## GLACIOLOGY

## Antarctic ice melt may speed up

Antarctica's vast ice sheets may be more vulnerable to warming than was thought.

Using a three-dimensional computer model, David Pollard of Pennsylvania State University in University Park and his colleagues identified two new ways in which ice sheets can collapse. Meltwater and rainfall can drain into crevasses in the ice, leading to vertical fractures. Moreover, heavy ice near the top of the sheets can break apart, shearing off huge chunks.

The results may help to explain how the East Antarctic Ice Sheet could have collapsed enough to cause the high sea levels that occurred during warm periods over the past 25 million years. Accounting for these mechanisms, the authors suggest that the West Antarctic Ice Sheet could collapse faster than predicted — over decades rather than centuries to millennia.

*Earth Planet. Sci. Lett.* 412, 112–121 (2015)

## MICROBIOLOGY

## Tuberculosis has history in its DNA

Key historical events such as the First World War drove the global spread of a strain of tuberculosis-causing bacteria that is prone to becoming resistant to drugs.

Thierry Wirth of the National Museum of Natural History in Paris and his colleagues collected 4,987 samples of the Beijing strain of *Mycobacterium tuberculosis*, isolated from patients from 99 countries, and analysed the microbe's DNA to trace its ancestry. They found that the strain originated in East Asia 6,600 years ago with the rise of agriculture. From there, it spread throughout the

## SOCIAL SELECTION

Popular articles on social media

## Ranking universities by happiness

Universities are often ranked using metrics for research income and academic impact, but scores such as those published by *Times Higher Education* do not say much about the quality of life of researchers at those institutions. A blog post ([go.nature.com/4bxozj](http://go.nature.com/4bxozj)) by structural biologist Jenny Martin that calls for new researcher-friendly metrics for ranking universities — including a happiness index — is drawing enthusiastic reviews on social media. "I am so going to work on increasing our Happiness-index in the group!" tweeted Jodie Bradby, a physicist at the Australian National University in Canberra. The proposal was "a new spin on

➔ **NATURE.COM**  
For more on popular papers:  
[go.nature.com/qulfg7](http://go.nature.com/qulfg7)

academic metrics ... that I can get behind", said Stephanie Januchowski-Hartley, a conservation scientist at Texas A&M University in Corpus Christi, also on Twitter.



world, increasing in prevalence when the human population grew in the nineteenth century, as well as when people were vulnerable to infection during the First World War and after HIV began to spread as epidemics. The drug-resistant strains that now affect Asia appeared when the Soviet Union — and its health system — collapsed in the 1990s.

*Nature Genetics* <http://dx.doi.org/10.1038/ng.3195> (2015)

## ECOLOGY

## Gold-rush threat to tropical forests

Deforestation due to gold mining is increasing in South America, particularly around biodiversity hotspots.

Gold mining has become more feasible in remote tropical forests owing to the drastic rise in demand and price for the metal over the

past decade. To study its impact, Nora Alvarez-Berrios and Mitchell Aide at the University of Puerto Rico-Río Piedras in San Juan analysed satellite images of tropical forests in South America from 2001 to 2013. They found that roughly 1,700 square kilometres of forest had been cleared (**pictured**) and only 250 km<sup>2</sup> was regenerated in and around gold-mining sites.

Although forest loss from gold mining is small compared with that from agriculture, for instance, it is accelerating — unlike deforestation as a result of other land-use changes. Moreover, nearly one-third of the losses are occurring within 10 km of protected areas.

*Environ. Res. Lett.* 10, 014006 (2015)

➔ **NATURE.COM**  
For the latest research published by Nature visit:  
[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

## NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION

# A case for a university happiness ranking

A whimsical blog post about using quality-of-life measures to rank research institutions earns praise online.

Chris Woolston

15 January 2015

Universities are often ranked using metrics for research income and academic impact, but scores such as those published by the *Times Higher Education* don't say much about the quality of life of researchers at those institutions. A blog post by structural biologist Jenny Martin that calls for new researcher-friendly metrics for ranking universities — including a happiness index — is drawing enthusiastic reviews on social media. “I am so going to work on increasing our Happiness-index in the group!” tweeted Jodie Bradby, a physicist at the Australian National University in Canberra. Stephanie Januchowski-Hartley, a conservation scientist at Texas A&M University in Corpus Christi wrote on Twitter:





**SJanuchowski-Hartley**  
@ConnectedWaters



 Follow

A new spin on academic \*metrics\* from  
[@JennyMartin\\_UQ](#) that I can get behind:  
[wp.me/p4bETE-iA](http://wp.me/p4bETE-iA) - thanks for the  
heads-up [@MeganEvans](#)



 **WordPress.com**

**imagine there's new metrics (it's easy if you try)**  
By WordPress.com @wordpressdotcom

Academia has become obsessed with metrics. Institutions jostle for the “top” positions in international rankings, departments are evaluated nationally to identify the “best”, and individuals are li...

[View on web](#)

RETWEETS

5

FAVORITES

4



3:56 AM - 13 Jan 2015

Martin, at the University of Queensland in Brisbane, Australia, wrote a post on her Cubistcrystal blog criticizing the current ranking system. She says that universities at the top of the list attract more and more PhD students, leading to a researcher oversupply that creates extra competition for jobs, unnecessary pressure and, ultimately, disenchantment. She proposes other measures that, in her view, would provide a more realistic view of life at a research institution. The Happiness index would be calculated using surveys of workplace satisfaction and the number of days of leave taken

<http://211.103.242.218/GW2/s/com/nature/www/G.http/news/a-case-for-a-university-happiness-ranking-1.16730>

1/3

without prompting. The Fairness Index would track the ratio of average salaries for men and women in leadership positions, and the K-index would measure 'kid-friendly' attributes, such as parental-leave policies and access to on-campus childcare. Such indices, if ever put in place, "might drive new, perhaps more socially just, workplace behaviours", she writes.

Martin admits in her post that she may be a dreamer, but "the new metrics I proposed were meant to get us thinking about how we can change the status quo," she said in an interview.

The proposals seemed appealing — if not especially plausible — to Eve Emshwiller, a botanist at the University of Wisconsin–Madison, who tweeted:



In a follow-up interview, Emshwiller said that if universities started competing to improve their rankings with such metrics, it would "add a lot to quality of life for those of us who work in these institutions."

And improving quality of life would result in better scientists, says Bradby. "I believe that scientists, like everyone else, work much better when they are connected, valued and supported," she says. One of her strategies is to give monthly awards for small achievements by those in her lab group, such as if they submit a paper or get a tricky experiment to work.

Including these indices would shake up university rankings, says Januchowski-Hartley. Universities that excel on the basis of conventional measures generally "aren't going to score high in happiness", she says.

Phil Baty, editor of the *Times Higher Education* World University Rankings, acknowledges that its ratings don't consider factors such as workplace satisfaction or pay equality when scoring universities. Collecting usable data on such issues from institutions around the world would be a huge challenge, he says. Still, he adds that the *Times Higher Education*



often reports on quality-of-life aspects of academia. "We take these issues very seriously."

*Nature* **517**, 415 (22 January 2015) doi:10.1038/517415f

---

For the best commenting experience, please login or register as a user and agree to our Community Guidelines. You will be re-directed back to this page where you will see comments updating in real-time and have the ability to recommend comments to other users.

## Comments

[Subscribe to comments](#)

There are currently no comments.

## See other News & Comment articles from *Nature*

*Nature* ISSN 0028-0836 EISSN 1476-4687

© 2015 Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.  
partner of AGORA, HINARI, OARE, INASP, CrossRef and COUNTER

# SEVEN DAYS

The news in brief

## EVENTS

### Lost probe found

A British-built Mars probe that failed to phone home after an attempted landing in December 2003 has been spotted on the red planet. Images from NASA's Mars Reconnaissance Orbiter suggest that the Beagle 2 craft landed safely but failed to deploy crucial equipment. For more, see [go.nature.com/wxhzki](http://go.nature.com/wxhzki) and page 412.

### Obesity approval

The US Food and Drug Administration on 14 January approved the first medical device designed to treat obesity by targeting nerve signalling. The Maestro implant, developed by EnteroMedics of St Paul, Minnesota, stimulates the vagus nerve between the oesophagus and stomach, blocking hunger signals to the brain (see *Nature* **511**, 18; 2014). In clinical trials, the device caused obese people to lose 8.5% more weight than people implanted with a dummy device.

## RESEARCH

### Global risks

Extreme weather has been rated the second most likely serious danger to nations and their economies by the World Economic Forum's 2015 *Global Risks* report. The non-profit organization, which is based in Geneva, Switzerland, and is behind the annual meetings of international leaders in the Swiss village of Davos, based the report on the opinions of around 900 experts. Interstate conflict was rated the most likely global risk, with extreme weather just behind. Natural catastrophes came in at number six, with failure of climate-change adaptation at seven. The top two risks in terms of impact rather than

likelihood were water crises and the spread of infectious diseases.

### Anthropocene start

The first atomic bomb test, on 16 July 1945 in New Mexico, could serve as the official start date of the Anthropocene epoch — a proposed unit of geological time that marks human influence on the planet. In a paper published on 12 January, 26 members of a stratigraphic working group argue that radioactive elements spread by this and subsequent nuclear tests provide a sensible transition point from the current Holocene epoch (J. Zalasiewicz *et al.* *Quat. Int.* <http://doi.org/zjff>; 2015). The working group

plans to come up with a final recommendation on whether and how to define the Anthropocene in 2016. See [go.nature.com/serlrs](http://go.nature.com/serlrs) for more.

### Millions die early

Alcohol, tobacco and unhealthy diets are shortening the lives of 16 million people with non-communicable diseases every year, the World Health Organization (WHO) has warned. In 2012, 38 million deaths were due to non-communicable diseases, according to a WHO report released on 19 January, and 16 million of these were premature. Margaret Chan, the WHO's director-general, called for increased government action to tackle the problem.

## BUSINESS

### Deepwater spill size

Oil company BP is responsible for discharging 3.19 million barrels of oil into the Gulf of Mexico during the 2010 Deepwater Horizon spill, a US federal court ruled on 15 January. The court says that the firm, headquartered in London, spilled around 4 million barrels, with 0.81 million later cleaned up — a larger net amount than BP's own estimate of the spillage, but less than that estimated by the US government. This ruling caps BP's potential fine for the disaster at US\$13.7 billion. Meanwhile, on 13 January, the US Environmental Protection Agency proposed revised rules for using chemical dispersants



ZOOM DOSSO/AFP/GETTY

## Progress in Ebola fight

Mali's government declared the Ebola epidemic over in its country on 18 January. The ministry of health said that there had been no confirmed cases of the virus in Mali since 6 December. Efforts continue to stop the spread of the disease

in the nations worst affected: Liberia (pictured), Guinea and Sierra Leone. In Liberia, the Ministry of Health and Social Welfare reported on 19 January that 12 of 15 counties in the country had had no new cases in the preceding week.

on oil spills, requiring improved studies of the chemicals' safety and efficacy.

## Nature merger

The publisher of *Nature* is to merge with Berlin-based publisher Springer Science+Business Media, their parent companies announced on 15 January. London-based Macmillan Science and Education, which owns all the Nature-branded journals, will combine its publishing portfolio with Springer's to create a firm with some 13,000 employees and a turnover of around €1.5 billion (US\$1.7 billion), the companies said. The deal was agreed by Holtzbrinck Publishing Group of Stuttgart, Germany, which owns Macmillan, and the London-based private equity firm BC Partners, which in 2013 bought Springer in a deal worth around €3.1 billion. See [go.nature.com/5rcn9i](http://go.nature.com/5rcn9i) for more.

### POLICY

## US methane rules

US oil and gas companies will have to reduce methane emissions by 40–45% below 2012 levels by 2025 under new rules announced by US President Barack Obama on 14 January. Methane emissions from fossil-fuel extraction (pictured) and distribution



contribute around 3% of total US greenhouse-gas emissions. The regulations build on a larger Obama-administration initiative to curb greenhouse-gas release that began last year and has already targeted car and power-plant emissions.

## GM-crop deal

The European Parliament has passed into law a political compromise that should allow more genetically modified (GM) crops to be cultivated in the European Union (EU). Under an agreement hashed out in December (see *Nature* <http://doi.org/xmq; 2014>), and passed on 13 January, crops will still be approved at the level of the EU, but member states will be able to ban their use locally. A deadlock between nations for and against genetic modification of crops has meant that hardly any GM crops are grown in the

EU. Some researchers have welcomed the compromise, although anti-GM groups and the biotechnology industry have been critical.

## Stimulus backlash

Researchers have complained that a mooted billion-euro investment fund designed to stimulate economies in the European Union (EU) will be partly financed by money that has been stripped from research budgets. On 13 January, the EU confirmed that some of the stimulus funds could come from cuts to the Horizon 2020 research-funding programme. The European University Association in Brussels expressed "strong disappointment" at the news.

## Clinical-trial data

All results of a clinical trial should be released within 18 months of the study's

completion, according to recommendations in a report from the US Institute of Medicine released on 14 January. The report also recommends that pharmaceutical and device companies that sponsor trials release at least a summary of the results and adverse effects within a year. Medical researchers have been pushing for greater openness of trial data, but some companies worry that releasing data too early could compromise trade secrets.

### PRIZES

## Crafoord winners

Two scientists who did pioneering work on genetic diversity and evolution in the 1960s and 1970s were awarded the 2015 Crafoord Prize, the Royal Swedish Academy of Sciences announced on 15 January. Richard Lewontin at Harvard University in Cambridge, Massachusetts, and Tomoko Ohta at the National Institute of Genetics in Mishima, Japan, will share the annual award, worth 6 million Swedish kronor (US\$735,000). The Crafoord Prize was established to recognize work that is not usually covered by the Nobel prizes, and it rotates between astronomy, geosciences and biosciences.

## Geology award

Volcanologist Stephen Sparks of the University of Bristol, UK, has won the US\$250,000 Vetlesen Prize, one of the most prestigious in the Earth sciences. The award, announced on 20 January, recognizes his field studies and laboratory work, which helped to standardize and modernize volcanology. Sparks led the UK government's scientific response to the volcanic eruption that began in 1995 on the Caribbean island of Montserrat. His work is now focused on studies of global volcanic risk.

► **NATURE.COM**

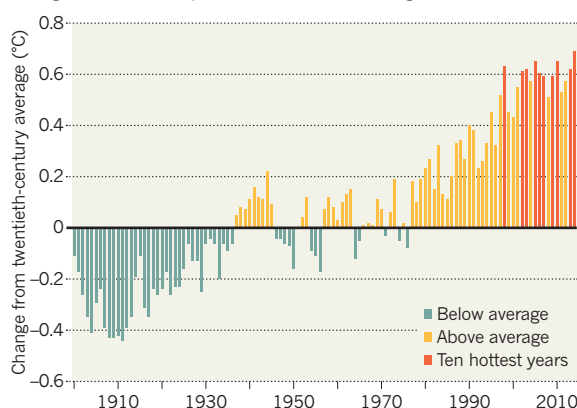
For daily news updates see:  
[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Last year was the hottest since records began in 1880, according to two separate analyses from researchers at NASA and at the US National Oceanic and Atmospheric Administration. Global surface temperatures for land and ocean were 0.69°C above the average for the twentieth century. Average surface temperatures have increased by 0.8°C since 1880. The record was reached in the absence of the El Niño ocean system, which often drives up temperatures (see [go.nature.com/ujnoif](http://go.nature.com/ujnoif)).

## 2014 TAKES TOP SPOT AS HOTTEST YEAR

Of the ten hottest years on record, nine have come since 2000. Global average and ocean temperatures both hit record highs in 2014.





# Attempts to predict terrorist attacks hit limits

Erratic human behaviour and incomplete information plague efforts to model risk.

Quirin Schiermeier

20 January 2015

[Print](#)

JOHN THYS/AFP/Getty Images

Police in Verviers, Belgium investigate after anti-terrorism operations on 15 January that left two suspects dead.

From France to Nigeria, the world is reeling from a wave of deadly terrorist attacks. As governments scramble to respond, researchers and analysts are attempting to calculate what the terrorism risk is and how it is changing. Such insights are crucial for those trying to foil attacks and for companies that offer insurance against terrorism, but it is neither an easy nor an exact science.

To calculate the risk of attacks, analysts must have access to a system for tracking events that have already happened. That is not straightforward, in part because terrorism is hard to define — the boundary between crude acts of terrorism and violent insurgencies is blurry.

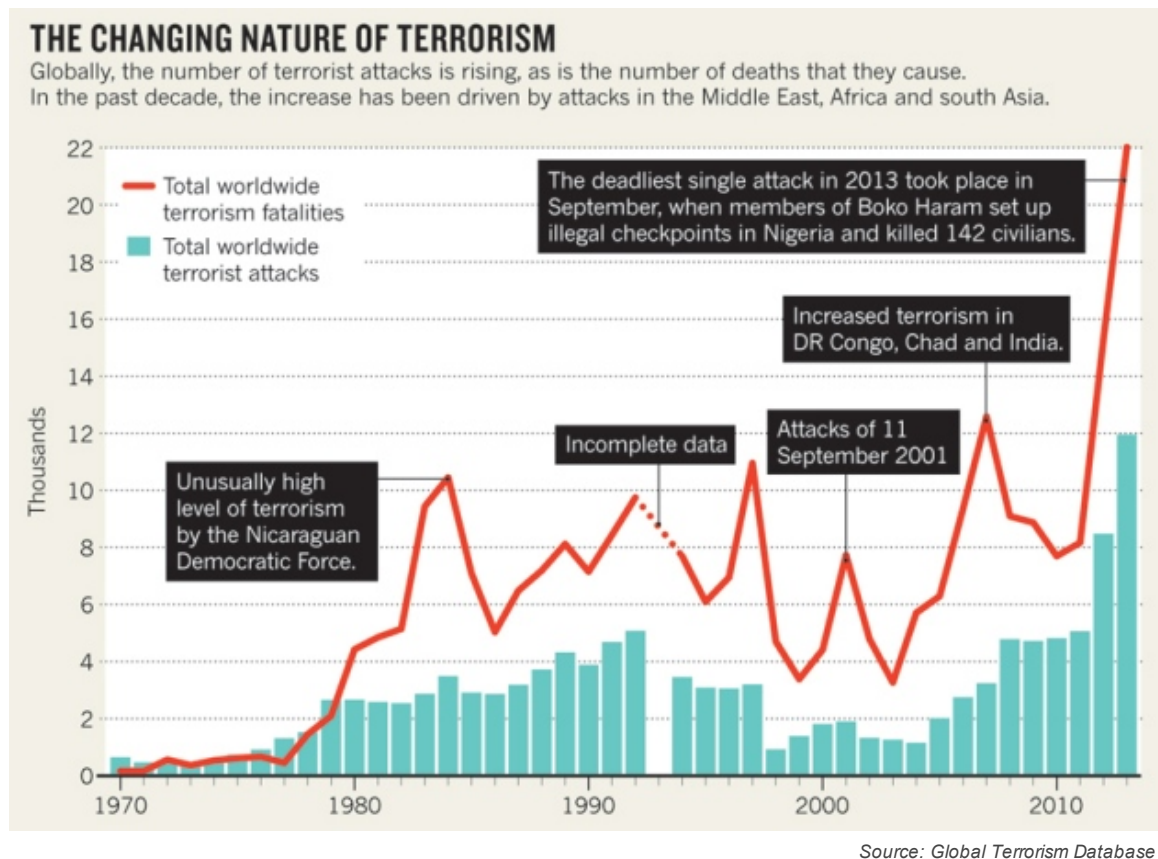
At the University of Maryland in College Park, Erin Miller oversees an effort to record terrorist events and make the data available to researchers. The group has settled on a definition of terrorism: “the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation”. The resulting Global Terrorism Database (GTD) provides insights into how terrorism has changed over the past few decades (see ‘The changing nature of terrorism’).

Overall, it shows that the number of reported events — from targeted assassinations to

**Top picks**  
from **nature** news

- Al Gore’s dream spacecraft gears up for launch
- Hidden hurdle for women in science
- Science and satire: The Paris attacks

hijackings and suicide bombings — has sharply increased globally since the turn of the millennium, mostly driven by attacks in the Middle East, Africa and south Asia. In the past ten years, half of all global terrorist attacks, and 60% of all fatalities, occurred in just three countries — Iraq, Afghanistan and Pakistan. By contrast, after the attacks in the United States on 11 September 2001, Europe and North America suffered relatively few successful attacks.



Counterterrorism measures vary between regions. In areas widely controlled by militant organizations such as the Afghan Taliban, ISIS or Boko Haram — responsible for an attack in Baga, Nigeria, this month that killed an estimated 150 people, among other events — counterterrorism forces are ineffective, corrupt or simply absent. However, many Western countries stepped up prevention measures in the wake of 9/11.

Plots in the West that were not foiled — including fatal attacks in Madrid in 2004, London in 2005, Norway in 2011 and most recently against the magazine *Charlie Hebdo* and in a supermarket in Paris — may be due to terrorist organizations switching to using fewer operatives, relying instead on small cells and lone actors, says Miller. This makes it less likely that details of a planned attack will leak out. “Al-Qaeda propaganda is clearly encouraging individuals to take up small attacks on their own,” she says.

As well as tracking attacks, the GTD is also being used to make predictions. Relying in part on the database, Aaron Clauset, a computer scientist at the University of Colorado Boulder, found in 2013 that the likelihood and magnitude of terrorist attacks are related by a phenomenon known as a power law. Smaller strikes with relatively few fatalities, such as in Paris, are sooner or later followed by a rare event with extremely high severity, such as 9/11 (A. Clauset and R. Woodard *Ann. Appl. Stat.* **7**, 1838–1865; 2013).

Power laws are a source of mathematical fascination because they describe a wide range of phenomena, including earthquakes and stock-market collapses. But Clauset used the relationship to predict that, given the historic rate of small-scale attacks, there is at least a 30% chance of another attack on the scale of 9/11 in the next decade somewhere in the world. He also points out that very large events are more likely under

#### Related stories

- Effective risk response needs a prepared mindset

a power law than they would be under a bell-shaped distribution — something that counterterrorism authorities should consider.

His predictions are relevant on a global scale, but they do not help to assess the risk to specific cities or companies, or the absolute magnitude of attacks. Besides counterterrorism agencies, one group tasked with addressing those issues is insurance underwriters, who must assess the terrorism cover that is mandatory for large businesses in the United States and a number of European countries.

- Psychologists seek roots of terror
- Looking for the roots of terrorism

**More related stories**

The 9/11 attacks caused insured losses exceeding US\$40 billion; a nuclear detonation over Chicago, Illinois, or Los Angeles, California, would cause hundreds of thousands of fatalities and financial losses far in excess of \$100 billion, according to Risk Management Solutions in Newark, California, which develops terrorism risk models for insurers.

However, such estimates are uncertain. The modelling resembles a game of poker in that it depends on erratic human behaviour and incomplete information. “You cannot reliably track terrorism risk without access to classified information,” says Erwann Michel-Kerjan, executive director of the Wharton Risk Center at the University of Pennsylvania in Philadelphia.

Because insurance companies have not had to pay out substantial claims since 9/11, terror premiums have steadily decreased. However, the terrorism landscape is in flux, and Michel-Kerjan says that governments, scientists and insurance companies are “pretty lost” when it comes to quantifying such a volatile risk. “*Charlie Hebdo* could have happened anywhere, any time,” he says. “We may not need to raise insurance premiums right away — but we will all need to pay more attention to what is happening in the Middle East, connect the dots, and do the science.”

*Nature* **517**, 419–420 (22 January 2015) doi:10.1038/517419a

See *World View* page 413 and *News* page 420

## Related stories and links

### From nature.com

- **Effective risk response needs a prepared mindset**  
20 January 2015
- **Psychologists seek roots of terror**  
20 January 2015
- **Looking for the roots of terrorism**  
15 January 2015
- **Science and satire**  
13 January 2015
- **How terror-proof is your economy?**  
15 October 2014
- **Arab liberals must stay in the game**  
08 August 2012
- **Arab liberals must stay in the game**  
08 August 2012
- **Is terrorism the next format for war?**  
12 July 2005

### From elsewhere



# Psychologists seek roots of terror

Studies raise prospect of intervention in the radicalization process.

Sara Reardon

20 January 2015



Print

Jason Athanasiadis/Corbis

Common themes are beginning to emerge from the study of terrorists' origins.

Anthropologist Scott Atran studies terrorists, seeking to understand what drives people to join groups such as the Islamist terrorist organization ISIS. The answers are sometimes surprising. One French Muslim extremist who sought to blow up an embassy traced his radicalization to a childhood incident: his sister bumped into a man on a Paris street and the man spat on the ground and called her a “dirty Arab”. “That’s when I knew what I was going to become,” the terrorist told Atran, a researcher at the University of Michigan in Ann Arbor.

Social scientists and law-enforcement agencies have long sought to understand the roots of terrorism, regardless of the ideology behind it. But interest has heightened in recent years, mostly over concerns about the rise of Islamic extremism. Common themes are beginning to emerge from such studies, giving researchers and policy-makers hope of developing effective methods to intervene in the radicalization process — and perhaps even to rehabilitate convicted or would-be terrorists.

Most budding extremists are motivated by disillusionment with society or a set of beliefs, and a desire to find others who share their values. But whether a disaffected person turns to a cult, gang or extreme ideology such as white supremacy or jihad depends on his or her environment and social networks. “On the surface, it may make no sense to compare a group like the IRA with al-Qaeda,” says psychologist John Horgan of the University of Massachusetts Lowell. “Only when we sit down do we find they have far more in common.”

## Top picks from **nature** news

- Al Gore's dream spacecraft gears up for launch
- Hidden hurdle for women in science
- Science and satire: The Paris attacks

With this in mind, a team led by Noemie Bouhana, a criminologist at University College London, is developing a computer model of the processes underlying radicalization and subsequent terrorist attacks. For example, the researchers hope to identify flaws in law-enforcement techniques or targets that are vulnerable to terrorism. Eventually, such models might be able to predict when the risk of an attack is high, Bouhana says. But models alone will never be able to pinpoint when a specific incident will occur, she says; that is probably impossible without information from intelligence agencies.

Even identifying which individuals might become terrorists is a difficult task. For instance, millions of people support the kind of militant Islam espoused by organizations such as al-Qaeda, but only a small percentage would be willing to kill for it, says Atran. Two studies that he published last month suggest that extremism arises, in part, when membership in a group reinforces deeply held ideals, and an individual's identity merges with the group's (S. Atran *et al. Cliodynamics* **5**, 41–57; 2014; S. Atran *et al. Proc. Natl Acad. Sci. USA* **111**, 17702–17703; 2014). “They can be low-lives, but once they lock into these values it doesn't matter, because they become heroic warriors,” says Atran.

His team interviewed a random sample of 260 people in two communities in Morocco that have produced an unusually high number of terrorists — including five of the main plotters of the 2004 Madrid train bombings, which killed 191 people. Many residents said that they wanted to fight to establish an ISIS-backed caliphate, a kind of Islamic government, in Syria. Those who believed in sharia, the Islamic code of law, said that they were willing to sacrifice for it in some way, but the degree of potential sacrifice increased dramatically if an individual had fused with a group with the same beliefs. In extreme cases, residents were willing to use violence or to allow their children to die to uphold their values.

Several dozen centres around the world are attempting to rehabilitate terrorists, with mixed results; one centre in Saudi Arabia claims an 88% success rate. A better understanding of radicalization could improve efforts to counteract it, Bouhana says. For instance, treatment centres might bring in former terrorists to talk about why they abandoned their causes.

Still, terrorists often leave extreme groups on their own. “They realize it's not the *Call of Duty*-style adventure they were led to believe,” Horgan says, referring to the computer combat game. Burnout can lead a person to disengage from the group, although not necessarily to give up radical beliefs. Others stop believing but do not leave their groups out of fear of retaliation.

Policy-makers have taken note of such research, and are increasingly recruiting social scientists as counterterrorism advisers. But Horgan says that there is still much more work to do. “We're only beginning to figure out what the right questions are,” he says. “Psychology's potential for the study of terrorism has yet to be realized.”

*Nature* **517**, 420–421 (22 January 2015) doi:10.1038/517420a

See *World View* page 413 and *News* page 419

#### Related stories

- Effective risk response needs a prepared mindset
- Attempts to predict terrorist attacks hit limits
- Looking for the roots of terrorism

#### More related stories

## Related stories and links

### From nature.com

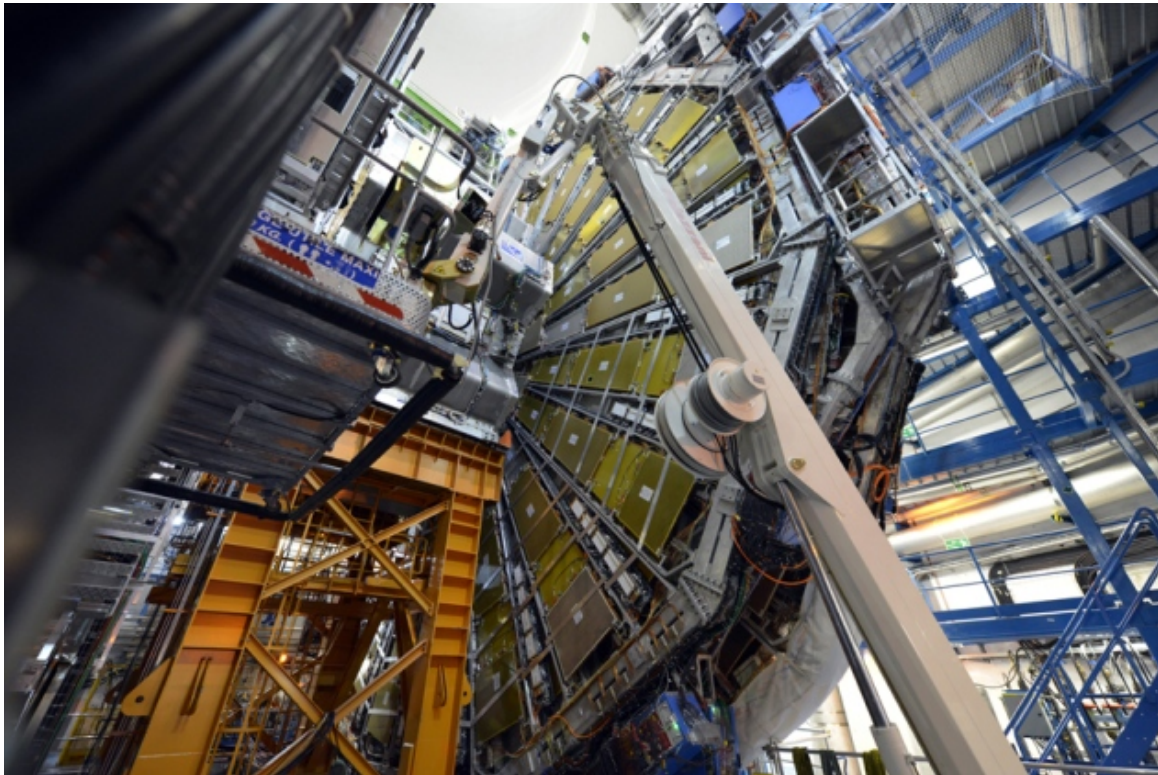
- **Effective risk response needs a prepared mindset**  
20 January 2015
- **Attempts to predict terrorist attacks hit limits**  
20 January 2015
- **Looking for the roots of terrorism**

# Crunch time for pet theory on dark matter

Thought to make up the Universe's missing matter, WIMPs are running out of places to hide.

Davide Castelvecchi

21 January 2015

[Print](#)

Harold Cunningham/Getty Images

Scientists hope to create WIMPs when the Large Hadron Collider comes back online.

After decades of searching and multiple no-shows, it is crunch time for a leading theory of what comprises dark matter, the mysterious stuff thought to make up around 85% of the Universe's matter.

The Large Hadron Collider (LHC) at CERN, Europe's particle-physics lab near Geneva, Switzerland, is scheduled to restart in March after a major upgrade. It is widely seen as the last chance in a generation to create — and thus confirm — theoretical particles known as WIMPs, or weakly interacting massive particles. A super-sensitive 'direct-detection' experiment, which is designed to catch naturally occurring WIMPs streaming from the heavens, is also due to start this year.

At the same time, the failure so far to glimpse WIMPs at either the LHC or through direct-detection experiments, combined with surprise signals from others, is fuelling suggestions that dark matter is made of something else. A range of alternatives that were previously considered underdog candidates now look "less exotic", says Kevork Abazajian, a theorist who studies particle cosmology at the University of California, Irvine.

Whatever dark matter is, most astronomers believe that it is real. The amount of ordinary, 'visible' matter does not produce enough gravity to explain the speed at which stars move inside galaxies, or at which galaxies move inside galaxy clusters. Dark matter would solve this mystery because it does not absorb or scatter light, making its presence known only

## Top picks from **nature** news

- Al Gore's dream spacecraft gears up for launch
- Hidden hurdle for women in science
- Science and satire: The Paris attacks



by its gravitational pull on normal matter. Different theories posit different suggestions as to what kind of particle would have these properties.

WIMPs are theoreticians' darlings. They are relatively heavy — somewhere between 1 gigaelectronvolt, or roughly the mass of one proton, and 1 teraelectronvolt — and thus would be relatively slow, or 'cold'. These properties fit well with the current best models for the evolution of the Universe, in which haloes of cold dark matter are the prime movers in the formation of galaxies and galaxy clusters. WIMPs also fix problems in two separate branches of physics: particle physics and cosmology. The mass of the WIMP and the strength of its interactions with other particles would help to explain why the Higgs boson has the mass it does. But these figures also mean that WIMPs would have been synthesized at just the right rate in the early Universe for the creation of the abundances that theory requires today — a coincidence that is dubbed the 'WIMP miracle'.

Yet despite being on physicists' most-wanted list, WIMPs remain on the run. When the LHC shut down for maintenance in 2013, its WIMP searches had come up empty. And the most sensitive of the direct-detection searches, carried out by the Large Underground Xenon (LUX) experiment at the Sanford Underground Research Facility in Lead, South Dakota, found no WIMPs during its first major run in 2013 (see *Nature* <http://doi.org/zh8>; 2013).

But WIMPs could show up this year. The souped-up LHC will smash together protons with combined energies of 13 teraelectronvolts — versus the 8 teraelectronvolts of the previous run: the extra energy will be capable of creating particles that were not possible to make before. Also, this summer, a WIMP detector called the XENON1T experiment, based under Gran Sasso in central Italy is scheduled to go live. It will have a sensitivity 50 times greater than LUX's, says XENON1T spokesperson Rafael Lang, a particle physicist at Columbia University in New York.

But the string of disappointments means that some theorists are already beginning to back away from WIMPs and look at alternatives, says cosmologist Scott Dodelson of the Fermi National Accelerator Laboratory in Batavia, Illinois.

One possibility is neutrinos, particles that interact weakly with other types of matter, a requirement for dark matter. The properties of the three known types of neutrino are not quite right for dark matter. But a hypothetical fourth type — called the 'sterile' neutrino, because it would interact even more weakly than its cousins — could be a suitable candidate.

In December 2014, the European Space Agency's Planck observatory released a map of ancient radiation from the early Universe that all but ruled out the existence of a sterile neutrino with a small mass like the other neutrinos, as did results from the Daya Bay neutrino experiment in China's Guangdong province.

But back in February 2014, astrophysicist Esra Bülbül of Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, and her collaborators reported a mysterious photon signal coming from 73 galaxy clusters (E. Bulbul *et al.* Preprint at <http://arxiv.org/abs/1402.2301>; 2014). The photons' wavelength was consistent with the decay of sterile neutrinos that weigh about 7 kiloelectronvolts — at least 30,000 times heavier than ordinary neutrinos.

Their results, posted on the arXiv repository, unleashed a flurry of proposed mechanisms by which dark matter could have produced the signal. Super-heavy neutrinos would still be light compared with WIMPs. According to conventional theories, this would make them 'warm' and they would not be a good fit with models of the Universe's evolution. But Abazajian says that the early Universe could have produced cold heavy neutrinos, and that these particles could fit with the current best models of galaxy formation, in some cases better than WIMPs (see, for example, K. N. Abazajian *Phys. Rev. Lett.* **112**,161303; 2014). "The WIMP miracle does have a theoretical nicety to it," says Abazajian — but, he adds, a heavy neutrino solves many of the same problems as WIMPs do.

#### Related stories

- European probe shoots down dark-matter claims
- Model Universe recreates evolution of the cosmos
- Physics: Broaden the search for dark matter

#### More related stories

Evidence for a heavy neutrino “would be really revolutionary”, Dodelson says. So far, however, teams that have attempted to reproduce Bülbül’s result have had mixed success.

Another WIMP alternative is the axion, a hypothetical particle proposed in the 1970s. Axions would spontaneously transmute into photons in a magnetic field, providing a means to detect them. Although a few experiments have failed, physicists led by Leslie Rosenberg of the University of Washington in Seattle is currently upgrading the sensitivity of its Axion Dark Matter Experiment. The team says that it will make or break the case for the axion.

Even Rosenberg has not given up on WIMPs, however. “There is nervousness about WIMP dark matter,” he says. But, “the LHC 2015 data set will need exploration before you should get too nervous.”

*Nature* **517**, 422–423 (22 January 2015) doi:10.1038/517422a

## Related stories and links

---

### From nature.com

- **European probe shoots down dark-matter claims**  
02 December 2014
- **Model Universe recreates evolution of the cosmos**  
07 May 2014
- **Physics: Broaden the search for dark matter**  
05 March 2014
- **Dark-matter search considers exotic possibilities**  
03 January 2014
- **No sign of dark matter in underground experiment**  
30 October 2013

### From elsewhere

- **XENON1T**
  - **LUX**
  - **Axion Dark Matter eXperiment**
  - **Planck**
  - **ATLAS results on supersymmetry**
  - **CMS results on supersymmetry**
  - **Daya Bay Reactor Neutrino Experiment**
- 

For the best commenting experience, please login or register as a user and agree to our Community Guidelines. You will be re-directed back to this page where you will see comments updating in real-time and have the ability to recommend comments to other users.

## Comments

**Subscribe to comments**

There are currently no comments.

**See other News & Comment articles from *Nature***

**Nature** ISSN 0028-0836 EISSN 1476-4687

with the decay of sterile neutrinos that weigh about 7 kiloelectronvolts — at least 30,000 times heavier than ordinary neutrinos.

Their results, posted on the arXiv repository, unleashed a flurry of proposed mechanisms by which dark matter could have produced the signal. Super-heavy neutrinos would still be light compared with WIMPs. According to conventional theories, this would make them ‘warm’ and they would not be a good fit with models of the Universe’s evolution. But Abazajian says that the early Universe could have produced cold heavy neutrinos, and that these particles could

fit with the current best models of galaxy formation, in some cases better than WIMPs (see, for example, K. N. Abazajian *Phys. Rev. Lett.* **112**, 161303; 2014). “The WIMP miracle does have a theoretical nicety to it,” says Abazajian — but, he adds, a heavy neutrino solves many of the same problems as WIMPs do.

Evidence for a heavy neutrino “would be really revolutionary”, Dodelson says. So far, however, teams that have attempted to reproduce Bülbül’s result have had mixed success.

Another WIMP alternative is the axion, a hypothetical particle proposed in the 1970s.

Axions would spontaneously transmute into photons in a magnetic field, providing a means to detect them. Although a few experiments have failed, physicists led by Leslie Rosenberg of the University of Washington in Seattle is currently upgrading the sensitivity of its Axion Dark Matter Experiment. The team says that it will make or break the case for the axion.

Even Rosenberg has not given up on WIMPs, however. “There is nervousness about WIMP dark matter,” he says. But, “the LHC 2015 data set will need exploration before you should get too nervous.” ■

## SYNTHETIC BIOLOGY

# Safety boost for GM organisms

*Engineered microbes kept in check with a synthetic building block.*

BY ELIE DOLGIN

Critics of genetic engineering have long worried about the risk of modified organisms escaping into the environment. A biological-containment strategy described this week in *Nature*<sup>1,2</sup> has the potential to put some of those fears to rest and to pave the way for greater use of engineered organisms in areas such as agriculture, medicine and environmental clean-up.

Two US teams have produced genetically modified (GM) bacteria that depend on a protein building block — an amino acid — that does not occur in nature. The bacteria thrive in the laboratory, growing robustly as long as the unnatural amino acid is included in their diet. But several experiments involving 100 billion or more cells and lasting up to 20 days did not reveal a single microbe capable of surviving in the absence of the artificial supplement.

“Our strains, to the extent that we can test them, won’t escape,” says Dan Mandell, a synthetic biologist at Harvard Medical School in Boston, Massachusetts, and an author on one of the two studies describing the strategy.

The microbes also do not swap their engineered DNA with natural counterparts because they no longer speak life’s shared biochemical language. “Establishing safety and security from the get-go will really enable broad and open use of engineered organisms,” says Farren Isaacs, a synthetic biologist at Yale University in New Haven, Connecticut, who led the other study.

Biocontainment could provide added safety in the biological production of drugs or fuels, where microbes can be kept separate from their surroundings. But the modified bacteria could also permit controlled release into the human body or the environment. “Containment might

no longer be of the physical kind,” says Tom Ellis, a synthetic biologist at Imperial College London who was not involved in the research.

The new technique originated in the laboratory of George Church, a geneticist at Harvard Medical School. Two years ago, Church and his team (which included Isaacs) reported the synthesis of a strain of *Escherichia coli* that had a reprogrammed genetic code<sup>3</sup>. Instead of recognizing a particular DNA triplet known as the amber stop codon as an order to terminate protein synthesis, the recoded bacterium read the same instruction as a directive to incorporate a new kind of amino acid into its proteins.

**“What we’re now starting to talk about is a really, completely synthetic organism.”**

Church and Isaacs have independently made this engineered microbe reliant on unnatural amino acids. The Isaacs team used genomic sequencing to identify sites in essential bacterial proteins where the microbes could incorporate synthetic amino acids without affecting overall function, whereas Church’s group started with the protein structures and added elements to help integrate and accommodate the artificial amino acids.

“This is really the culmination of a decade of work,” says Church.

These organisms are also more resistant to viruses than their natural counterparts because of the mismatch between the genetic code of the virus and that of its host<sup>3</sup>. Looking ahead, Church and his team are working to co-opt seven different codons, instead of just one. “That would be more than enough to be resistant to all viruses

and to create a lot of opportunity for safety,” Church says.

Isaacs has also developed a different safeguarding system, in which *E. coli* can grow only in environments containing synthetic chemicals needed for gene expression. He described the work this month in *Nucleic Acids Research*<sup>4</sup>. Another research team led by Jef Boeke at the New York University Langone Medical Center and Patrick Yizhi Cai at the University of Edinburgh, UK, has been working on a similar strategy in yeast. Commonly used in industry and biotechnology, yeast has its genetic material packaged in chromosomes similarly to animals and plants rather than bacteria.

“That’s a strategy that is going to be more easily adaptable to other organisms beyond *E. coli*,” says Isaacs. His team is now engineering a bacterium that is dependent on synthetic chemicals as well as on artificial protein building blocks. “I think ultimate solutions for robust biocontainment will involve multiple approaches that are deployed at the same time in a single organism,” he says.

Such a beast will present a real challenge for regulators, says Todd Kuiken, senior research associate for the Science and Technology Innovation Program at the Woodrow Wilson International Center for Scholars in Washington DC. “What we’re now starting to talk about is a really, completely synthetic organism,” Kuiken says. “How do you evaluate that once you put it out into the field?” ■

**NATURE.COM**  
To hear more about making safer GM organisms, visit:  
[go.nature.com/jrn2t](http://go.nature.com/jrn2t)

1. Mandell, D. J. *et al.* *Nature* <http://dx.doi.org/10.1038/nature14121> (2015).
2. Rovner, A. J. *et al.* *Nature* <http://dx.doi.org/10.1038/nature14095> (2015).
3. Lajoie, M. J. *et al.* *Science* **342**, 357–360 (2013).
4. Gallagher, R. R., Patel, J. R., Interiano, A. L., Rovner, A. J. & Isaacs, F. J. *Nucleic Acids Res.* <http://dx.doi.org/10.1093/nar/gku1378> (2015).

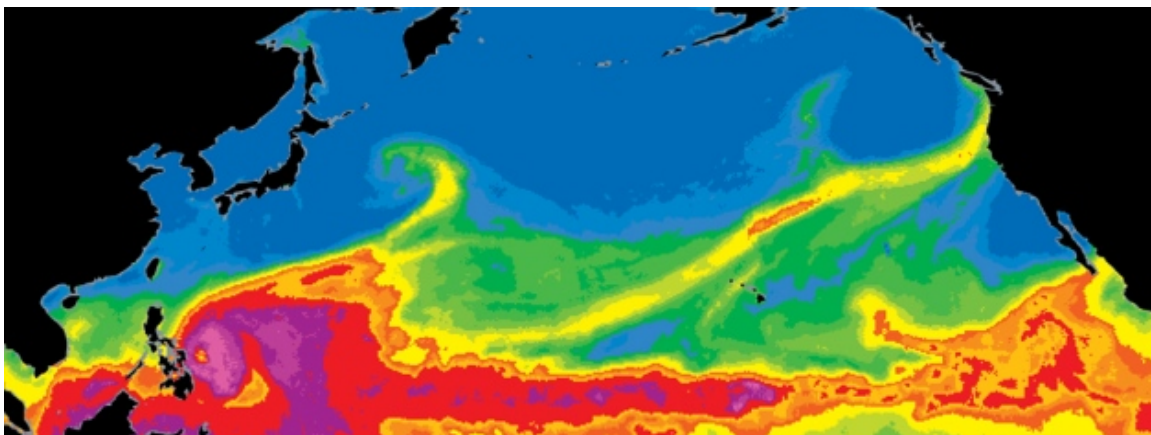


# CalWater 2015 targets atmospheric rivers off California

Meteorologists investigate airborne jets that bring both floods and drought relief.

Alexandra Witze

20 January 2015


[Print](#)

Darren Jackson/NOAA ESRL

A satellite image showing water-vapour concentration reveals an atmospheric river (yellow) streaming northeast across the Pacific Ocean.

Californians call it the Pineapple Express: a weather pattern that zips across the Pacific Ocean from Hawaii, delivering not baskets of tropical fruit, but buckets of rain and snow.

In meteorological terms, the Pineapple Express is an atmospheric river, a narrow band of air that carries huge amounts of moisture. For the next six weeks, meteorologists will be plying the eastern Pacific by air and sea, in the hope of catching several atmospheric rivers barrelling towards the coast. It is the biggest push yet to understand these phenomena, which have received serious scientific attention only in the past decade.

Atmospheric rivers get their start over warm tropical waters; they then flow eastwards and towards the poles a kilometre or two above the ocean surface. They may stretch for thousands of kilometres, but are only a few hundred kilometres wide. When they hit land, they start to drop their moisture in torrential downpours or blizzards.

“When we have too many atmospheric rivers, floods can occur, and when we don’t have enough we gradually fall into drought,” says Marty Ralph, a meteorologist at the Scripps Institution of Oceanography in La Jolla, California, and a leader of the field campaign.

In Europe, atmospheric rivers affect mostly the western part of the continent, but they can be felt as far inland as Poland. In North America, the entire west coast is affected, and parts of the central and eastern United States occasionally feel the effects of atmospheric rivers that develop over the Gulf of Mexico. The moisture is often welcome, bringing up to half of the year’s water supply in affected areas<sup>1</sup>. A 2013 study found that as many as three-quarters of all droughts in the Pacific Northwest between 1950 and 2010 had been brought to an end by atmospheric-river storms<sup>2</sup>. California has been stricken by drought for years (see *Nature* **512**, 121–122; 2014), but last month, an atmospheric river dropped enough rain to erase one-third of the water deficit of one major reservoir in just two days.

**Top picks**  
from **nature** news

- Al Gore’s dream spacecraft gears up for launch
- Hidden hurdle for women in science
- Science and satire: The Paris attacks

Climate change may bring stronger and more frequent atmospheric rivers, because the warmer the atmosphere is, the more water it can hold, says David Lavers, a meteorologist at Scripps who is not involved in the project. “The more you know about how the atmosphere behaves,” he says, “the better position you’re in to prepare for extreme events.”

The current US\$10-million field campaign, called CalWater 2015, is a massive push to capture the physics of atmospheric rivers as they make landfall. A research vessel from the US National Oceanographic and Atmospheric Administration (NOAA), the *Ronald H. Brown*, is sitting several hundred kilometres off the California coast, waiting to position itself beneath an atmospheric river whenever one forms. The ship carries a suite of instruments operated by the US Department of Energy (DOE) that will peer upwards to study aerosol concentrations and other properties of the atmosphere once a river has formed.

Four aeroplanes — two belonging to NOAA, and one each to the DOE and NASA — will fly directly into the atmospheric river. They will release dropsondes, miniature parachutes carrying an array of instruments that record weather data as they descend.

One major goal is to work out how aerosols affect atmospheric rivers, says Kimberly Prather, an atmospheric chemist at Scripps and another leader of the project. A smaller field campaign in California between 2009 and 2011 suggested that contrary to expectations, aerosols served as nuclei for ice crystals to grow within atmospheric rivers and later produce snow<sup>3</sup>. “It’s Mother Nature’s way of seeding things,” says Prather.

CalWater 2015 will measure how aerosol particles interact with atmospheric rivers, both offshore and when they reach the coast. Supercomputer simulations will examine how airborne dust affects the amount and type of precipitation that eventually falls, says Ruby Leung, an atmospheric modeller with the Pacific Northwest National Laboratory in Richland, Washington.

That information should help water managers to predict what atmospheric rivers may bring. Reservoir engineers in northern California typically release precious water from their reservoirs in the winter, so as to have enough space behind the dam to cope with the threat of flooding. With better knowledge of when atmospheric rivers might arrive and how much water they might carry, says Ralph, engineers should be able to manage that winter release more effectively.

*Nature* **517**, 424–425 (22 January 2015) doi:10.1038/517424a

## References

1. Lavers, D. A. & Villarini, G. *J. Hydrol.* <http://dx.doi.org/10.1016/j.jhydrol.2014.12.010> (2014).

Show context

2. Dettinger, M. D. *J. Hydrometeorol.* **14**, 1721–1732 (2013).

Show context

Article ISI

3. Creamean, J. M. *et al. Science* **339**, 1572–1578 (2013).

Show context

Article PubMed ISI ChemPort

## Related stories and links

### From nature.com

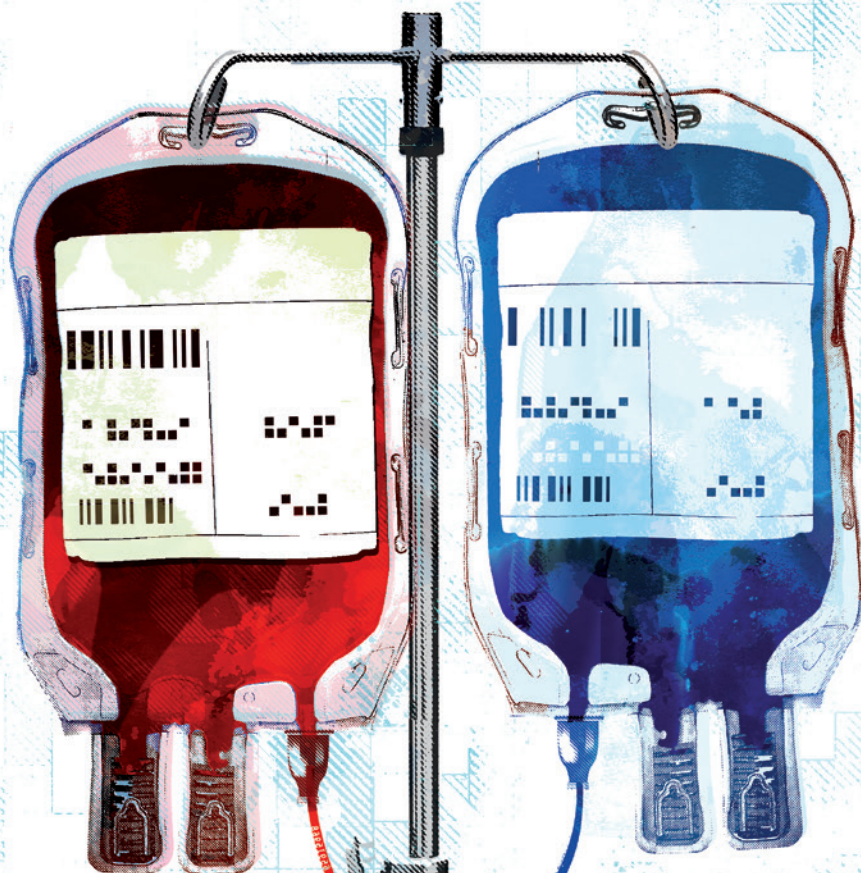
- **Major cloud-seeding test gives mixed results**

11 December 2014

### Related stories

- Major cloud-seeding test gives mixed results
- Native ecosystems blitzed by drought
- Climate change makes extreme weather more likely to hit UK

### More related stories



# Blood to blood

By splicing animals together, scientists have shown that young blood rejuvenates old tissues. Now, they are testing whether it works for humans.

BY MEGAN SCUDELLARI

Two mice perch side by side, nibbling a food pellet. As one turns to the left, it becomes clear that food is not all that they share — their front and back legs have been cinched together, and a neat row of sutures runs the length of their bodies, connecting their skin. Under the skin, however, the animals are joined in another, more profound way: they are pumping each other's blood.

Parabiosis is a 150-year-old surgical technique that unites the vasculature of two living animals. (The word comes from the



Greek *para*, meaning ‘alongside’, and *bios*, meaning ‘life’.) It mimics natural instances of shared blood supply, such as in conjoined twins or animals that share a placenta in the womb.

In the lab, parabiosis presents a rare opportunity to test what circulating factors in the blood of one animal do when they enter another animal. Experiments with parabiotic rodent pairs have led to breakthroughs in endocrinology, tumour biology and immunology, but most of those discoveries occurred more than 35 years ago. For reasons that are not entirely clear, the technique fell out of favour after the 1970s.

In the past few years, however, a small number of labs have revived parabiosis, especially in the field of ageing research. By joining the circulatory system of an old mouse to that of a young mouse, scientists have produced some remarkable results. In the heart, brain, muscles and almost every other tissue examined, the blood of young mice seems to bring new life to ageing organs, making old mice stronger, smarter and healthier. It even makes their fur shinier. Now these labs have begun to identify the components of young blood that are responsible for these changes. And last September, a clinical trial in California became the first to start testing the benefits of young blood in older people with Alzheimer’s disease.

“I think it is rejuvenation,” says Tony Wyss-Coray, a neurologist at Stanford University in California who founded a company that is running the trial. “We are restarting the ageing clock.”

Many of his colleagues are more cautious about making such claims. “We’re not de-ageing animals,” says Amy Wagers, a stem-cell researcher at Harvard University in Cambridge, Massachusetts, who has identified a muscle-rejuvenating factor in young mouse blood. Wagers argues that such factors are not turning old tissues into young ones, but are instead helping them to repair damage. “We’re restoring function to tissues.”

She emphasizes that no one has convincingly shown that young blood lengthens lives, and there is no promise that it will. Still, she says that young blood, or factors from it, may hold promise for helping elderly people to heal after surgery, or treating diseases of ageing.

“It’s very provocative,” says Mark Mattson, chief of the Laboratory of Neurosciences at the US National Institute on Aging in Bethesda, Maryland, who has not been involved in the parabiosis work. “It makes you think. Maybe I should bank some blood of my daughter’s son, so if I start to have any cognitive problems, I’ll have some help,” he says, only half-joking.

## THE POWER OF TWO

Physiologist Paul Bert performed the earliest recorded parabiosis experiment in 1864, when he removed a strip of skin from the flanks of two albino rats, then stitched the animals together in hopes of creating a shared circulatory system<sup>1</sup>. Biology did the rest: natural wound-healing processes joined the animals’ circulatory systems as capillaries regrew at the intersection. Bert found that fluid injected into a vein of one rat passed easily into the other, work that won him an award from the French Academy of Sciences in 1866.

Since Bert’s initial experiments, the procedure has not changed much. It has been performed on hydra — small freshwater invertebrates related to jellyfish — frogs and insects, but it works best on rodents, which recover well from the surgery. Up to the mid-twentieth century, scientists used parabiotic pairs of mice or rats to study a variety of phenomena. For example, one team ruled out the idea that dental cavities are the result of sugar in the blood by using a pair of parabiosed rats, of which only one was fed a daily diet of glucose. The rats had similar blood glucose levels owing to their shared circulation, yet only the rat that

actually ate the sugar developed cavities<sup>2</sup>.

Clive McCay, a biochemist and gerontologist at Cornell University in Ithaca, New York, was the first to apply parabiosis to the study of ageing. In 1956, his team joined 69 pairs of rats, almost all of differing ages<sup>3</sup>. The linked rats included a 1.5-month-old paired with a 16-month-old — the equivalent of pairing a 5-year-old human with a 47-year-old. It was not a pretty experiment. “If two rats are not adjusted to each other, one will chew the head of the other until it is destroyed,” the authors wrote in one description of their work<sup>4</sup>. And of the 69 pairs, 11 died from a mysterious condition termed parabiotic disease, which occurs approximately one to two weeks after partners are joined, and may be a form of tissue rejection.

*“I thought, ‘Hey wait, they’re sharing blood. This could answer the question we’ve been asking for years.’”*

Today, parabiosis is performed carefully to reduce animal discomfort and mortality. “We observe the mice at length and have long discussions with our animal-care committee,” says Thomas Rando, a Stanford neurologist who has used the procedure. “We don’t take this lightly.” Mice of the same sex and size are socialized with each other for two weeks before attachment, and the surgery itself is done in a sterile setting with anaesthesia, heating pads and antibiotics to prevent infection. Using inbred lab mice, genetically matched to one another, seems to reduce the risk of parabiotic disease. Joined mice eat, drink and behave normally — and they can be separated successfully.

In McCay’s first parabiotic ageing experiment, after old and young rats were joined for 9–18 months, the older animals’ bones became similar in weight and density to the bones of their younger counterparts<sup>5</sup>. More than 15 years later, in 1972, two researchers at the University of California studied the lifespans of old–young rat pairs. Older partners lived for four to five months longer than controls, suggesting for the first time that circulation of young blood might affect longevity<sup>6</sup>.

Despite these intriguing findings, parabiosis fell out of use. Those who have studied the technique’s history speculate that researchers thought they had learned all they could from it, or that the bar for getting institutional approval for parabiosis studies had become too high. Whatever the reason, the experiments stopped. That is, until a stem-cell biologist named Irving Weissman brought parabiosis back to life.

## BACK TO THE SOURCE

Weissman learned to join mice together at the age of 16, under the supervision of a hospital pathologist in the small town of Great Falls, Montana, in 1955. His supervisor was studying transplantation antigens, proteins on the surface of transplanted cells or tissues that determine whether they are accepted or rejected by the host. Weissman remembers adding a fluorescent tracer to the blood of one mouse in a pair and watching it go back and forth between the animals. “It was really amazing,” he says.

He went on to spend three decades studying stem cells and regeneration in natural parabionts, sea squirts of the species *Botryllus schlosseri*. In 1999, Wagers, then a new postdoctoral fellow in Weissman’s Stanford lab, wanted to study the movement and fate of blood stem cells, so Weissman recommended that she use parabiotic mice and fluorescently label the cells she wanted

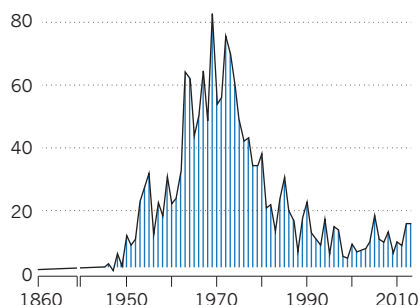
**➔ NATURE.COM**  
For a podcast on parabiosis and ageing, visit:  
[go.nature.com/berp8l](http://go.nature.com/berp8l)

# Share and share alike

Parabiotic experiments, in which two animals share a common bloodstream, were first attempted in the 1860s. By connecting animals with different qualities or conditions, scientists can investigate how blood factors, such as cells, proteins or hormones, influence health. In recent years, a few researchers have looked at heterochronic (old and young) mouse pairs to understand how young blood helps to repair many tissues.

## Publications on parabiosis

Parabiosis gained popularity during the 1960s and 1970s, but eventually fell out of wide practice.



to track in one animal of a pair. Wagers' experiments led to two rapid-fire discoveries on the nature and migration of blood stem cells<sup>7,8</sup>. It also inspired her Stanford neighbours.

In 2002, Irina Conboy, a postdoctoral fellow in Rando's lab, presented one of Wagers' papers at a journal-club meeting. Michael Conboy, Irina's husband and a postdoc in the same lab, was dozing in the back of the meeting room.

The mention of stitching mice together jolted him awake. "We had been in discussion for years that ageing seems to be all cells in the body, that all tissues seem to go to hell in a handbasket together," says Michael. Yet they had been unable to think of a realistic experiment with which to investigate what coordinates ageing throughout the body.

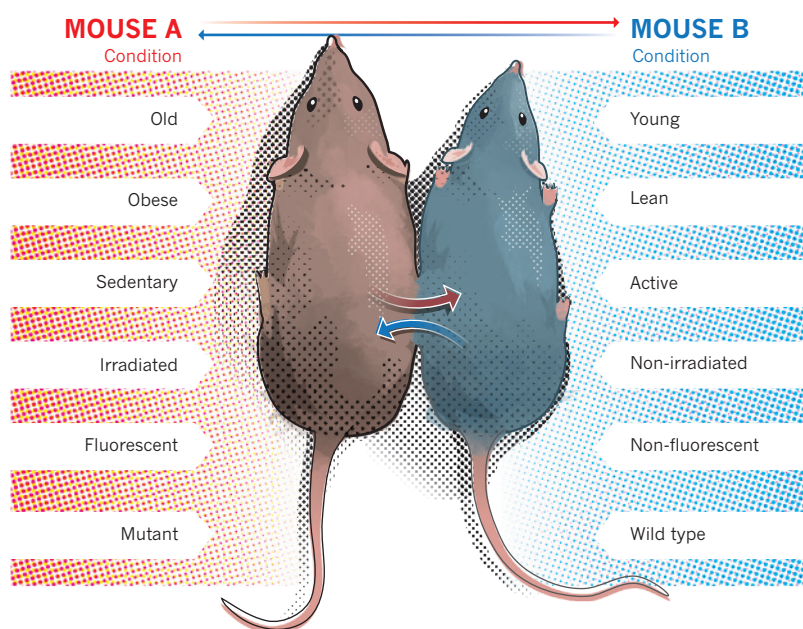
"I thought, 'Hey wait, they're sharing blood,'" says Michael. "This could answer that question we've been asking for years." At the end of the presentation, he ran up to Irina and Rando. He had not even finished his pitch before Rando said: "Let's do it."

The researchers teamed up with Wagers, who performed the old-young pairings for the experiment and taught Michael the technique (see 'Share and share alike'). Rando says that he did not expect the experiment to work, but it did. Within five weeks, the young blood restored muscle and liver cells in the older mice, notably by causing aged stem cells to start dividing again<sup>9</sup>. The team also found that young blood resulted in enhanced growth of brain cells in old mice, although the work was left out of their 2005 paper describing the results. All in all, the results suggested that blood contains the elusive factor or factors that coordinate ageing in different tissues.

After the team published its results, Rando's phone started ringing incessantly. Some of the calls were from men's health magazines looking for ways to build muscle; others were from people fascinated by the prospect of forestalling death. They wanted to know whether young blood extended lifespan. But despite the hints that this was true from the 1970s, no one has yet properly tested the idea. It would be an expensive, labour-intensive experiment.

## A simple surgery

A veterinary surgeon will anaesthetize the animals, peel away a thin layer of skin along their sides and stitch or staple the exposed surfaces together. Wound-healing processes join the bloodstreams through a capillary network, and in one to two weeks, the animals are pumping each other's blood.



Instead, members of the original research team branched out into separate efforts to determine what exactly in the blood is responsible for the rejuvenating effects. In 2008, Irina and Michael Conboy, by then at the University of California, Berkeley, linked<sup>10</sup> muscle rejuvenation to the activation of Notch signalling — which promotes cell division — or to the deactivation of the transforming growth factor (TGF)- $\beta$  pathway, which blocks cell division. Then, in 2014, they identified<sup>11</sup> one of the age-defying factors circulating in the blood: oxytocin, a hormone best known for its involvement in childbirth and bonding, and already a drug approved by the US Food and Drug Administration for inducing labour in pregnant women. Oxytocin levels decline with age in both men and women, and when injected systemically into older mice, the hormone quickly — within a couple of weeks — regenerates muscles by activating muscle stem cells.

## ALL THE ORGANS

Wagers was following up on the anti-ageing work at Harvard, where she had started her own lab in 2004. She recruited the help of experts in various organ systems to help her to evaluate the impact of young blood on their respective tissues. With neuroscientist Robin Franklin at the University of Cambridge, UK, her team showed<sup>12</sup> that young blood promotes repair of damaged spinal cords in older mice. With Harvard neuroscientist Lee Rubin, she found<sup>13</sup> that young blood sparks the formation of new neurons in the brain and olfactory system. And with cardiologist Richard Lee at Brigham and Women's Hospital in Boston, Massachusetts, she found<sup>14</sup> that it reverses age-related thickening of the walls of the heart.

With Lee, Wagers began screening for proteins that were particularly abundant in young blood but not old blood. One leapt out at them: growth differentiation factor 11, or GDF11. Wagers and Lee showed<sup>14</sup> that direct infusions of GDF11 alone were sufficient to physically increase the strength and stamina of muscles, as well as to reverse DNA damage inside muscle stem cells. No mouse

studies outside of Wagers lab have yet replicated the finding, but a similar protein in fruit flies extends lifespan and prevents muscular degeneration<sup>15</sup>.

It is perhaps fitting that parabiosis' newfound popularity has spread among labs with close ties. Wyss-Coray, who worked in the room next to Rando's lab, had previously discovered prominent changes in levels of proteins and growth factors in the blood of ageing humans and people with Alzheimer's disease. Following up on Rando's unpublished brain results, he used old-young mouse pairs to show<sup>16</sup> that old mice exposed to young blood did indeed have increased neuron growth, and that young mice exposed to old blood had reduced growth. Plasma alone had the same effects. "We didn't have to exchange the whole blood," says Wyss-Coray. "It acts like a drug." Next, the team looked at overall changes in the brain, and found that young plasma activates brain plasticity and memory formation in older mice, and increases learning and memory. "We could not believe that this worked," says Wyss-Coray.

Neither could the reviewers. The first time Wyss-Coray submitted the work to a journal, it was rejected, he says, responding that it was too good to be true. So his team spent a year repeating the experiments at the University of California, San Francisco — a different facility with different staff, instruments and tools. The researchers got the same results. "After that, I was really reassured," says Wyss-Coray. "I'm convinced it works."

His research, published last May<sup>17</sup>, caught the attention of a company in Hong Kong owned by a family with a history of Alzheimer's disease, which is characterized by neuron loss. One family member's condition had reportedly temporarily improved after they received a plasma transfusion. So the company put forward the initial funding to translate Wyss-Coray's approach to human clinical trials. Wyss-Coray formed a start-up company, Alkahest in Menlo Park, California, and in September 2014 it began a randomized, placebo-controlled, double-blind trial at Stanford, testing the safety and efficacy of using young plasma to treat Alzheimer's disease. Six out of a planned 18 people with Alzheimer's, all aged 50 or above, have already begun to receive plasma harvested from men aged 30 or younger. In addition to monitoring disease symptoms, the researchers are looking for changes in brain scans and blood biomarkers of the disease.

## BAD BLOOD?

Wagers is eager to see the results, but she worries that a failure would be difficult to interpret and so could set the whole field back. Plasma from a 30-year-old donor may not contain factors beneficial to patients with Alzheimer's, for example. She, Rando and others would prefer to see testing for a specific blood factor or combination of known factors synthesized in the lab, for which the mechanism of action is fully understood.

There are also lingering concerns as to whether activating stem cells — which is what the young blood most often seems to do — over a long period of time would result in too much cell division. "My suspicion is that chronic treatments with anything — plasma, drugs — that rejuvenate cells in old animals is going to lead to an increase in cancer," says Rando. "Even if we learn how to make cells young, it's something we'll want to do judiciously."

Michael Conboy is concerned for another reason: he has seen enough paired mice die of parabiotic disease to be cautious about trying it in humans. "I would be leery" of any trial in which significant amounts of blood or plasma were transfused into an older person regularly, he says. Alkahest's chief executive, Karoly Nikolich, says that he understands the safety concerns, but he emphasizes that millions of blood and plasma transfusions have been carried out safely in humans.

The initial Alkahest study is expected to conclude by the end of this year, and the company plans to initiate further studies testing

young plasma in the treatment of different types of dementia and age-related conditions.

All the caution over young blood is justified, given the history of dashed hopes in the anti-ageing field. In the past two decades, researchers have identified the anti-ageing properties of numerous treatments, including calorie-restricted diets; resveratrol, a chemical found in the skin of grapes; telomerase, an enzyme that protects the integrity of chromosomes (see Books & Arts,

*"You often have these lucrative markets emerge on a slender foundation of credible work."*

page 436); rapamycin, an immune-suppressing drug that extends lifespan in mice; and stem cells, which decline in function and number as people age.

Only two of these — caloric restriction and rapamycin — have been shown to reliably slow or reverse the effects of ageing across many mammalian tissue types, but neither has turned into an anti-ageing treatment. The former has produced conflicting results in primates; the latter has toxic side effects.

Young blood, by contrast, seems to turn back the effects of ageing, potentially with few known safety concerns in humans and, so far, with corroborated results from parabiotic ageing studies in multiple labs. But scientists and ethicists still worry about the treatment being tried in people outside approved clinical trials before evidence on its safety and effectiveness is in. Unlicensed stem-cell transplants are already a booming industry, warns Mattson, and unlicensed transfusion of young blood would be even easier.

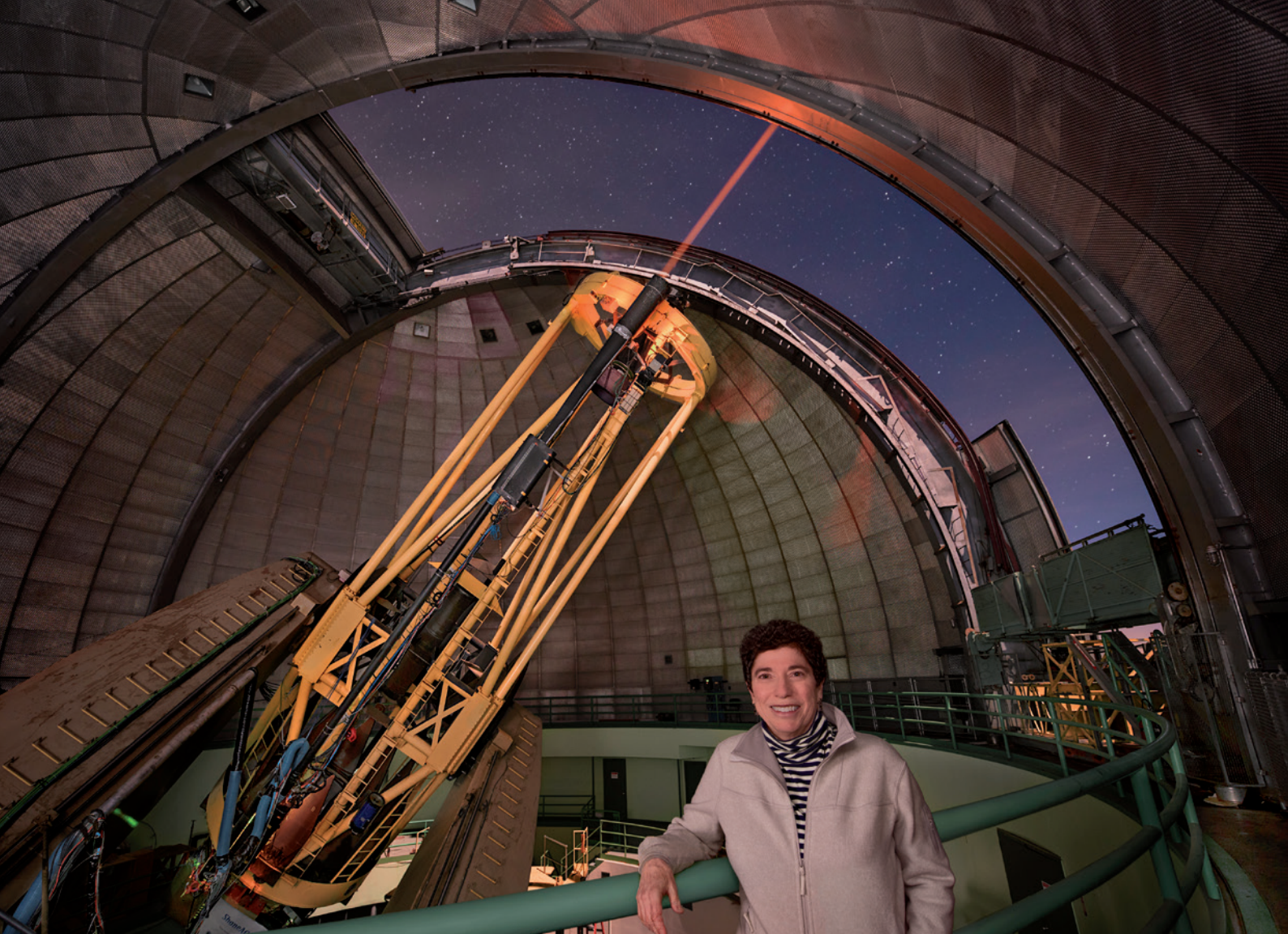
"You often have these lucrative markets emerge on a slender foundation of credible work," says Leigh Turner, a bioethicist at the University of Minnesota in Minneapolis who has studied the anti-ageing field.

For now, any claims that young blood or plasma will extend lifespan are false: the data are just not there. An experiment to test such claims would take upwards of six years — first waiting for the mice to age, then for them to die naturally, then analysing the data. "If we had funding to do this, I'd do it. But we don't," says Michael Conboy. Still, he adds, "I hope that someone, somewhere is." ■

*Megan Scudellari is a science journalist based in Boston, Massachusetts.*

1. Bert, P. J. *Anatomie Physiologie* **1**, 69–87 (1864).
2. Kamrin, B. B. *J. Dent. Res.* **33**, 824–829 (1954).
3. McCay, C. M., Pope, F., Lunsford, W., Sperling, G. & Sambhavaphol, P. *Gerontologia* **1**, 7–17 (1957).
4. McCay, C. M., Pope, F. & Lunsford, W. *Bull. New York Acad. Med.* **32**, 91–101 (1956).
5. Horrington, E. M., Pope, F., Lunsford, W. & McCay, C. M. *Gerontologia* **4**, 21–31 (1960).
6. Ludwig, F. C. & Elashoff, R. M. *Trans New York Acad. Sci.* **34**, 582–587 (1972).
7. Wright, D. E., Wagers, A. J., Gulati, A. P., Johnson, F. L. & Weissman, I. L. *Science* **294**, 1933–1936 (2001).
8. Wagers, A. J., Sherwood, R. I., Christensen, J. L. & Weissman, I. L. *Science* **297**, 2256–2259 (2002).
9. Conboy, I. M. *et al. Nature* **433**, 760–764 (2005).
10. Carlson, M. E., Hsu, M. & Conboy, I. M. *Nature* **454**, 528–532 (2008).
11. Elabd, C. *et al. Nature Commun.* **5**, 4082 (2014).
12. Ruckh, J. M. *et al. Cell Stem Cell* **10**, 96–103 (2012).
13. Katsimpardi, L. *et al. Science* **344**, 630–634 (2014).
14. Loffredo, F. S. *et al. Cell* **153**, 828–839 (2013).
15. Demontis, F., Patel V. K., Swindell, W. R. & Perrimon, N. *Cell Rep.* **7**, 1481–1494 (2014).
16. Villeda, S. A. *et al. Nature* **477**, 90–94 (2011).
17. Villeda, S. A. *et al. Nature Med.* **20**, 659–663 (2014).





# LASER FOCUS

*By firing lasers into the sky, Claire Max has transformed the capabilities of current — and future — telescopes.*

BY ANN FINKBEINER

On clear, moonless evenings, most of the biggest optical telescopes around the world begin the night's observations by firing a golden laser beam at the sky.

Claire Max does not like to take credit for this astronomical light show, even though the lasers' widespread use is a tribute to her three-decade campaign to perfect and promote them — an effort that was recognized on 16 January when the American Astronomical Society awarded her its 2015 instrumentation prize. For Max, an astronomer at the University of California, Santa Cruz, self-aggrandizement would be — unbusinesslike. And she is all business; even her way of speaking is careful, like someone who feels obliged to stand behind every word she says. Her passion is reserved for the technology itself. "I still get gripped by it," she says, showing off photograph after photograph of telescopes, lasers and thin beams of light shining upwards as straight as a ruler.

The lasers, Max explains, are a crucial element of the telescopes' adaptive optics, which correct for turbulence in the atmosphere. Without adaptive optics, stars and galaxies viewed at high magnification will dance, distort and blur like stones seen at the bottom of a stream. With adaptive optics, they will remain steady and sharp, allowing telescopes on the ground to routinely equal or exceed the clarity obtained by NASA's Hubble Space Telescope. This capability has allowed current-generation telescopes to carry out high-resolution studies of objects ranging from moons in the outer Solar System to stars at the centre of the Milky Way. And now it is enabling the construction of telescopes measuring 20–40 metres across, as much as four times the diameter and 16 times the light-gathering power of any now in existence.

Max has been involved in this development from its early days: from the first demonstration of laser-assisted adaptive optics to building the prototype and then establishing a centre that spread the technology to telescopes around the world.

Yet Max's greatest triumph has also become her greatest challenge. Last October, at an age when other astronomers might be looking forward to retirement, the 68-year-old Max agreed to serve as interim director of the University of California Observatories (UCO) — the organization responsible for all the astronomical hardware owned by one of the biggest state university systems in the United States. And in that role, 'interim' or not, Max finds herself navigating the professional and cultural chaos in astronomy being triggered by the cost of these next-generation behemoths.

There are three of these telescopes in various stages of planning and construction, each with a price tag in the order of US\$1 billion. That cost, says Max, poses a quandary for their owners and funders — among

LAURIE HATCH



Claire Max stands next to the 3-metre telescope at California's Lick Observatory.

essential workhorses for individual researchers and training grounds for young astronomers? Or should they fight to find creative ways to keep all the doors open?

Max's instinct is to fight — using her unique combination of warmth, empathy and determination. So far, she is winning. After three decades of persuasion and consensus-building in pursuit of adaptive optics, says Andrea Ghez, an astronomer at the University of California, Los Angeles, Max has developed a sure instinct for making connections among engineers, academics, funding officers, university administrators and all the others who have a say in telescope decisions.

These are powerful players, says Ghez — “gorillas at the table who'd like you for lunch”. And to deal with them, she says, you need someone like Max: “a gorilla with finesse”.

## FIRST LIGHT

Max took her first look through a telescope in the early 1950s, when the Manhattan native was just 8 years old. “And that was it,” she says. “I can still close my eyes and see the mountains of the Moon. So I became an astronomy nut.”

That passion led her first to Princeton University in New Jersey, where she earned a PhD in astronomy; then to a postdoctoral appointment at the University of California, Berkeley; to a staff position at the Lawrence Livermore National Laboratory in California; and in 1983, to membership in the Jasons: a group of scientists who meet from time to time to give technical advice on national security, often for the US Department of Defense.

The first Jason study she joined was motivated by the US Air Force's desire to identify potentially hostile satellites — a task for which the atmosphere was just as big a barrier as it was for astronomers. Even with clear skies and a good telescope, turbulence smears out details smaller than about 1 arcsecond in angular diameter — good enough to look up at the Hubble telescope, which is similar in size and altitude to spy satellites, and tell that it is a cylinder, but not much else.

Astronomers had already come up with a potential solution: a flexible mirror that could reflect the light coming into the telescope and deform under computer control. In principle, the distortions introduced by the mirror would exactly cancel out those produced by the atmosphere, restoring the image to near-perfection. But first, the distortion had to be measured, preferably by looking at what the atmosphere did to a bright ‘guide’ star near the target. And bright stars were not always available near the fast-moving targets of interest to the Pentagon.

That is why the Air Force had asked the Jasons for help — which Max and her colleagues provided in a classified report. Just shine a laser upwards along the telescope's axis, they said. If the laser were tuned to the correct wavelength, the beam would then encounter a naturally occurring layer of sodium atoms floating in the atmosphere about 90 kilometres up and cause the sodium to fluoresce, producing a bright-yellow spot visible from the ground — in effect, a guide star available anywhere in the sky (see ‘Untwinkle the stars’).

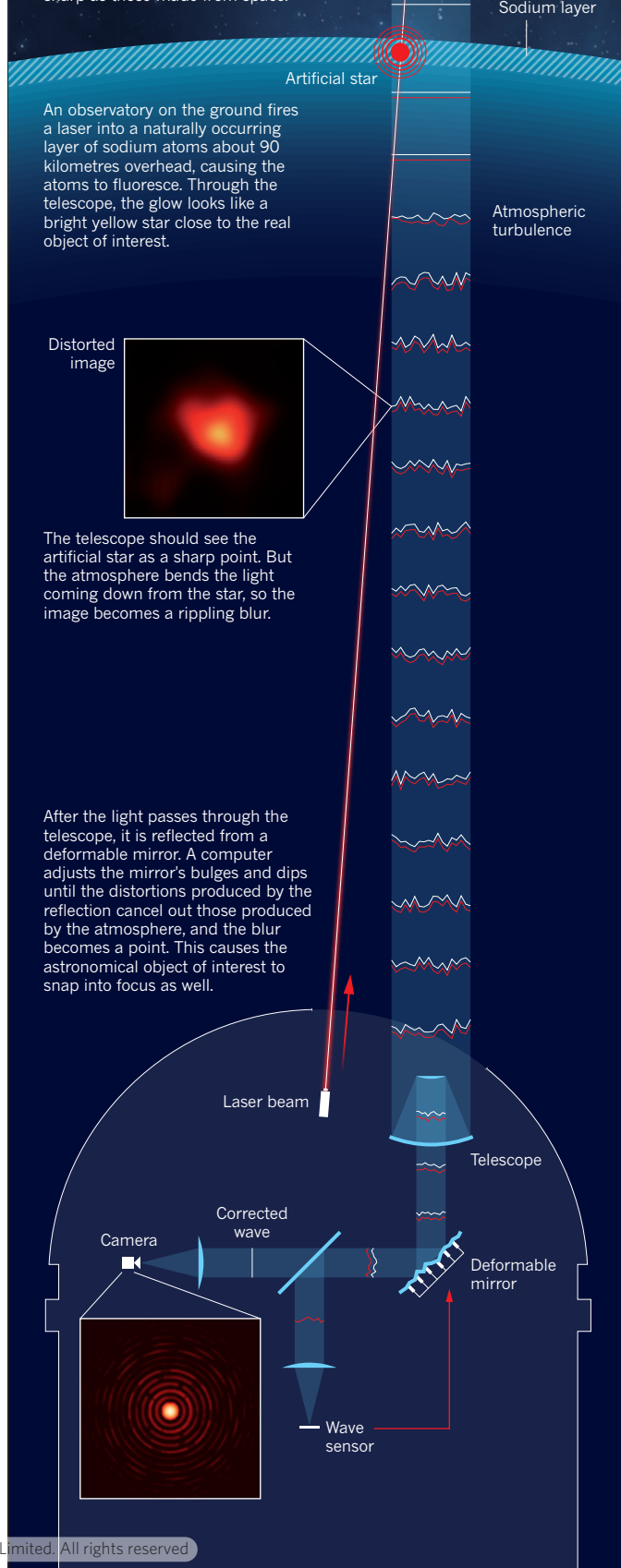
But Max went further. Knowing that the sodium laser guide star would also be invaluable for astronomers, she came up with an additional design that the Air Force had not asked for, but was better adapted to research needs. “I thought it was so cool,” she says.

Unfortunately, that design was classified along with all the other sodium laser guide-star technology. So Max, together with sympathetic Air Force scientists who knew the proper channels to follow, spent the next seven years lobbying the military to take the wraps off. And they won: in May 1991, Air Force scientist Robert Fugate was allowed to describe the artificial guide star at an open meeting of the American Astronomical Society in Seattle, Washington. Reactions were mixed:

them the UCO, a key partner in the Thirty Meter Telescope (TMT) that started construction last year atop Mauna Kea in Hawaii. How do they pay for all their older, smaller telescopes? Should the owners give in to financial pressure and close the facilities — even though the telescopes are still

## UNTWINKLE THE STARS

Laser guide-star technology can measure and reverse the distortion caused by turbulence in the atmosphere, allowing ground-based telescopes to produce images as sharp as those made from space.



astronomers who had been working on similar technology but did not have the money to get very far were annoyed to learn that the military had already built it in secret. But most were just excited at the prospect of getting blur-free images anywhere they wanted. “It was just pandemonium,” recalls Fugate. They asked “a million questions”.

For all the astronomers’ enthusiasm, however, the system was still technically demanding, expensive and in need of development that universities could not afford. “People were writing about it,” Max says, “but they weren’t putting it on telescopes.” So shortly after the declassification, Max decided that astronomers needed a proof-of-principle. The idea for how to do that hit during lunch with a colleague, Herb Friedman. “We looked at each other,” she says, “and we said, ‘Well crap, we’re from Livermore, we do lasers.’”

Indeed, the laboratory had an enormous underground laser that was normally used to separate isotopes, but could be tuned to sodium wavelengths. Max and Friedman therefore arranged to have an access cover removed from the roof of the laser tunnel, and a mirror installed to bounce the horizontally pointing laser light through the hole into the sky. Then they set up a small telescope beside the beam to look for the artificial guide star and measure the atmospheric distortions. The set-up must have looked eerie in the darkness, says Max — the tent that covered the telescope glowed yellow with stray laser light; the beam itself could be seen for ten kilometres or more — and one local woman kept calling the police to report that a UFO was sucking up all Livermore’s secrets. But their demonstration did prove that the design would work the way that Max expected.

The next step was to master the complex optics and engineering required for the laser system to function on a real telescope. A prototype that Max and her colleagues deployed at the UCO’s Lick Observatory in San Jose in the mid-1990s eventually showed that, at least at longer wavelengths, the system allowed the observatory’s 3-metre telescope to reach the finest-possible resolution permitted by the wave nature of light (D. T. Gavel *et al. Proc. SPIE* **4007**, 63–70; 2000).

But even that failed to persuade astronomers to embrace the technology. At the time, people thought the laser guide star was “so complicated it would never run in harness” with other astronomical instruments, says Connie Rockosi, one of Max’s colleagues. So Max decided that the technology needed “a community of practice” — a central home in which users could learn how to build laser guide stars for themselves. That idea became the Center for Adaptive Optics, which was funded by a \$5 million, 10-year grant from the US National Science Foundation (NSF), and which opened on the Santa Cruz campus in 1999, with Max eventually as its head. By 2010, says Max, when the grant ended and the centre had to close its doors, it had grown from a handful of people to nearly 100.

That did the trick. The laser-assisted adaptive-optics systems have to be custom-built for each telescope and are still pricey, running to several million dollars apiece. But astronomers, many of whom were trained at Max’s centre, have now retrofitted the technology onto every optical telescope for which it makes sense. That includes almost all the flagship telescopes that currently rank as the largest in the world, from the twin Keck 10-metre telescopes atop Mauna Kea to the four identical 8.2-metre instruments comprising the European Very Large Telescope in Chile.

## GOOD RESOLUTION

The bigger the telescope, the more advantage it can get from adaptive optics. At 1 micrometre, a wavelength in the infrared part of the spectrum, which is particularly useful for astronomy, the Hubble Space Telescope’s 2.4-metre mirror can produce images with a resolution 0.11 arcseconds. At that wavelength, with the help of the laser guide star, the Lick 3-metre can do somewhat better: 0.08 arcseconds. But 8-metre instruments like those comprising the European quartet in Chile can get all the way down to 0.03 arcseconds — almost four times better than Hubble.

That kind of visual acuity has allowed astronomers to track the stars orbiting the Milky Way’s central black hole, image exoplanets around other suns, observe the common, low-mass stars known as brown dwarfs and pursue many other once-impossible studies. In these areas, says David Silva, director of the US National Optical Astronomy Observatory in Tucson, Arizona, “we couldn’t have made any advances from the ground” without adaptive optics.

The systems’ largest impact, however, will be on the 20–40-metre telescopes that are now under development: the European Extremely Large Telescope and the Giant Magellan Telescope, both in Chile, and the TMT on Mauna Kea. Telescopes this big can collect enough light to study faint, far-off objects such as the first galaxies to form after the Big Bang — but would hardly be worth their billion-dollar cost if they had a resolution of just 1 arcsecond.

Unfortunately, that price tag is also why there are only three of these megatelescopes — which means, in turn, that only a small fraction of astronomers will ever get to use them. At the same time, ‘have-not’ astronomers at institutions not affiliated with one of the large projects are losing access to the 3- and 4-metre telescopes — even though these smaller instruments are often ideal for large-scale surveys of the sky, or targeted observations of relatively nearby objects. Citing a flat budget

and its investment in large projects such as the ALMA radio telescope in Chile, the NSF has withdrawn support from optical telescopes in this class. “For the general astronomer, it’s harder and harder to get time,” said Richard Barvainis, programme officer for astronomy at the NSF. “It’s becoming a major issue.”

At the UCO, which is facing its own cash crunch, Max is now in charge of a perfect microcosm of the situation she helped to create. In September 2013, provost Aimée Dorr declared that the University of California’s support for the Lick Observatory would end within five years, thanks to construction costs on the TMT; maintenance expenses on the Keck telescopes; and the UCO’s declining budget, which has halved over the past decade. The astronomical community responded with a firestorm of protest. It insisted that Lick, in operation since 1888, is still valuable as a test bed for new instruments and as a training ground for graduate students.

So last autumn, Max began her term as interim director of the UCO by walking into Dorr’s office and saying “What can we do to make this a win-win?”. Dorr was willing to try: Max was not only “credible, straightforward, honest, and moderate”, she says, but she could create budgets that at least one previous director had said were impossible. Over the next several months, Max sorted through the UCO’s convoluted partnerships and budgets, found ways to move around various pots of money, calmed the firestorm and got the university’s funding for Lick reinstated.

Max’s next priority is to devise a comprehensive strategic plan for Lick, as well as for the two Keck telescopes. Because Lick has adaptive optics and it is relatively available to astronomers, it could be used for high-precision surveys of a few hundred to a few thousand objects such as quasars — studies, says Max, that “you can peck away one by one and add up to a survey with good statistics”.

Max is determined to have this plan in place when the university announces its choice for the permanent director of the UCO. That new director might be her — she has applied for the position — but if not, she says, she will simply go back to being the astronomer-engineer-community builder she has been all along. Along those lines, she is working on multi-laser adaptive-optics systems that can both correct for the atmosphere and widen the field of view. She is also training a particular brand of graduate student — the kind, she says, who “has one hand in instruments and one in astronomy”.

That breadth of skills has allowed Max to do what she has done, says Ghez. “A typical astronomer could never do it.” ■

**Ann Finkbeiner** is a freelance writer in Baltimore, Maryland.



# COMMENT

**TECHNOLOGY** Holding our algorithmic society to account **p.435**



**FILM** Documentary follows two scientists' quest for longevity **p.436**

**COMMUNICATION** Risk lessons from low-key forecast of Typhoon Haiyan **p.438**

**FUNDING** Polarized responses to debate over indirect costs **p.438**

STR/AFP/GETTY



A resident of China's Hubei province clears the Fuhe river of dead fish, thought to have been poisoned by high levels of ammonia.

## Four gaps in China's new environmental law

Implementation and accountability will remain challenging, especially at the local level, warn **Bo Zhang** and **Cong Cao**.

On 1 January, a new environmental protection law (EPL) took effect in China. It is the nation's first attempt to harmonize economic and social development with environmental protection.

The EPL is perceived as the most progressive and stringent law in the history of environmental protection in China. It

details harsher penalties for environmental offences — for example, for acts of tampering and falsifying data, discharging pollutants covertly and evading supervision. It contains provisions for tackling pollution, raising public awareness and protecting whistle-blowers. It places more responsibility and accountability on local governments and law-enforcement agencies

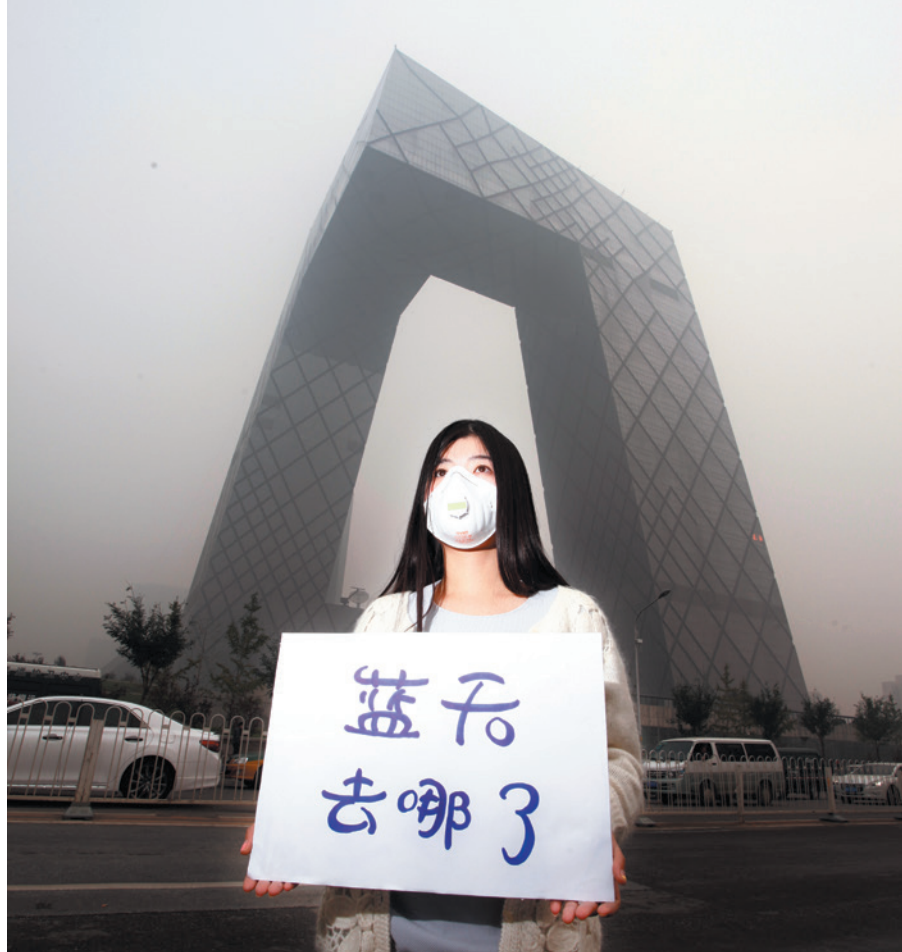
and sets higher standards for enterprises.

Yet the law is not enough. For the following reasons it will face many challenges<sup>1</sup>.

### FOUR GAPS

First, the power of the new law is limited. The EPL can be trumped by other legislation such as the specific agriculture, forestry, grassland and water laws, which ▶





'Where has the blue sky gone?' asks a citizen in Beijing during a period of heavy smog in 2014.

► are aimed at protecting these resources. Consequently, the departments responsible for managing these natural resources could challenge EPL provisions.

Many countries have enacted basic environmental-protection legislation to avoid this situation. For example, in 1969, the United States formulated the National Environmental Policy Act, which placed legal obligations and liabilities for environmental protection on all federal agencies<sup>2</sup>. In 1993, Japan replaced its Basic Law for Environmental Pollution and its Natural Environment Preservation Law with a comprehensive Basic Environment Law<sup>3</sup>.

Second, enforcement of the EPL will be hampered by the fragmented and overlapping structure of environmental governance in China. At present, the environment and natural resources are overseen by several agencies, including the Ministry of Environmental Protection, the ministries of water resources, land and resources, and agriculture, the State Oceanic Administration and the State Forestry Administration<sup>4</sup>. China's carbon-tax and emissions-trading systems, overseen by the National Development and Reform Commission, are not included in the law.

Other countries practise independent and unified environmental supervision. In the United States, the Council on Environmental Quality, which reports to the president, coordinates federal agencies and implements environmental policies at the federal, state and local government level. Meanwhile, the

Environmental Protection Agency is responsible for overall environmental governance representing the federal government and working through independent law enforcement<sup>3</sup>. Italy's Ministry of the Environment, Land and Sea takes overall responsibility for protecting the country's land, ocean, rivers, wetlands and forests. Other examples include Germany's Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety, and Brazil's Ministry of the Environment.

Third, despite increasingly damaging pollution in China, the new EPL fails to acknowledge citizens' basic right to an environment fit for life. So far, 149 developed and developing countries, including Russia, South Korea and the Philippines, have acknowledged in their constitutions a substantial right to environmental quality<sup>5</sup>. Although China's

new law clearly gives citizens, civic groups and non-governmental organizations (NGOs) the right to obtain environmental information and to participate in environmental governance, these parties (with the exception of NGOs that fulfil certain criteria) are not allowed to bring lawsuits against the government if there are, for example, serious lapses in air or water quality. A public environmental litigation system is needed to provide

**"The new law fails to acknowledge citizens' basic right to an environment fit for life."**

cost-effective redress for those affected.

Fourth, enforcement and implementation of the law may be foiled by a lack of capacity and by conflicts of interest. Environmental governance in China is mainly exercised locally by environmental protection bureaus. The environment ministry primarily provides guidance to these bodies, but does not have sufficient authority over them. Local governments (which are often more interested in economic growth) keep a firm grip on the staffing and financing of the environmental protection bureaus and hence on their decision-making<sup>4</sup>. This means that, in effect, the protection bureaus do not have the power to impose severe penalties on companies or individuals in breach of environmental statute, nor do the bureaus face legal or other consequences of failure to do so.

## PLUGGING THE GAPS

In October 2014, the Central Committee of the Communist Party of China decided to establish a recording, communication and accountability system. This is intended to limit judicial intervention by government officials, and to integrate promoting the rule of law into the remit of officials at all levels. The system could restrain officials, and indeed the party itself, from advising judges on how to decide on court cases, including those pertaining to the EPL.

Fundamentally, as Alex Wang, an environmental law specialist at the University of California, Los Angeles, has written, environmental protection in China needs to be raised to "a level of priority previously reserved only for the most important party-state mandates, such as economic growth, social stability, and the one-child policy"<sup>6</sup>.

Effective environmental governance needed a new law. Now it requires robust implementation mechanisms, accountability regimes and institutional arrangements. ■

**Bo Zhang** is a research fellow at the Information Center, Ministry of Environmental Protection, China.

**Cong Cao** is associate professor and reader at the School of Contemporary Chinese Studies, University of Nottingham, UK. e-mails: zhangbo@mep.gov.cn; cong.cao@nottingham.ac.uk

1. McElwee, C. R. *Environmental Law in China: Mitigating Risk and Ensuring Compliance* (Oxford Univ. Press, 2011).
2. Hahn, R. W. *Nat. Resour. J.* **34**, 305–348 (1994).
3. Ren, Y. *Int. Rev. Environ. Strategies* **1**, 79–96 (2000).
4. Economy, E. C. *The River Runs Black: The Environmental Challenge to China's Future* 2nd edn (Cornell Univ. Press, 2010).
5. Boyd, D. R. *The Status of Constitutional Protection for the Environment in Other Nations* (David Suzuki Foundation, 2014).
6. Wang, A. L. *Harvard Environ. Law Rev.* **37**, 365–440 (2013).

The views expressed are those of the authors and do not necessarily reflect the official view of China's Ministry of Environmental Protection.





## BIG DATA

# Stealth control

Steven Aftergood appraises a study on the hidden impacts of personal data collection.

Everyone who uses the Internet for entertainment, education, news or commerce is implicated in a web of data collection whose breadth surpasses ordinary awareness.

Last May, a US Senate investigation reported that a single visit to a popular tabloid-news website triggered activity on more than 350 other web servers. Most of those contacts, including delivery of advertisements, are likely to be benign. But they typically deposit a software 'cookie' on the visitor's computer; these enable the identification and tracking of visitors, generating digital profiles of their interests and patterns of online behaviour.

The Internet relies on user data to produce tailored advertising revenue that can support its growth and free use. But digital profiling ultimately helps to construct what law professor Frank Pasquale calls "the black box society". As his exposé of that name shows, this is a society in which basic functions are performed in deliberate obscurity through the collection and algorithmic manipulation of personal data.

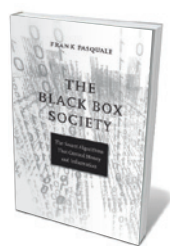
Black-box algorithms can be used to draw plausible inferences about a subject's location, age, medical condition, political

affiliation and so on. The US retail giant Target, for instance, used data about purchases of vitamin supplements and oversized bags to deduce — for marketing purposes — whether its customers were pregnant. Because applying such algorithms to specific tasks can be economically advantageous for advertisers or lenders, the algorithms typically remain undisclosed as 'trade secrets'. But Pasquale worries that they can be used to shape what we know, how we are perceived and what opportunities we will be afforded. Increasing reliance on Google or other search engines, he notes, fosters greater dependence on their operating principles of selection and prioritization — which are largely opaque, beyond the fact that they reflect (and reinforce) popularity. Information sources that search engines neglect or suppress will not be discoverable. And if a search engine's

algorithm were to subtly favour the company's own corporate interests, for instance, we might never know it.

Meanwhile, our reputations are defined in important ways by automated credit scoring and other algorithmic profiling activities used to determine, for instance, credit-worthiness or suitability for employment. Data from pharmacies concerning prescription purchases have been used by health insurers to deny individuals coverage. Yet our ability to inspect, correct or challenge such profiling is significantly limited. Black-box decisions also figure prominently in the finance industry, which uses them to exploit differential access to market-related information. Finance has acquired an undesired mystique, Pasquale believes, by adopting computationally intense procedures to model and anticipate market behaviour. By exaggerating the validity of such models and concealing their risks from investors and regulators, some Wall Street firms exacerbated the financial crisis of 2008.

The insulation of black-box practices from public inspection or evaluation is close to the root of the problem: it tends to preclude independent oversight, error correction and even free-market competition. Remarkably, the US Congress has found



**The Black Box Society: The Secret Algorithms That Control Money and Information**  
FRANK PASQUALE  
Harvard University Press: 2015.

LEE WOODGATE/GETTY



► it easier to elicit oversight information from intelligence fortresses such as the National Security Agency than from some Silicon Valley firms.

Pasquale provides an informative survey of developments in the representative fields of search, reputation and finance to bolster his argument that a laissez-faire approach to algorithmic decision-making is taking us to places where most of us will not want to go. As the power of advertising providers such as Google AdSense grows, for example, many online publications are seeing a decline in their advertising revenue. Homeland security 'fusion centres' are integrating government data collection (which is constrained by law) with unregulated information from private data brokers, in the name of information sharing. More promisingly, Pasquale points to the US Treasury's little-known Office of Financial Research, sometimes called "the CIA of finance", which aims to provide regulators with real-time intelligence on financial markets. The book is full of instructive anecdotes on such matters, backed by useful research.

There are occasional lapses. Pasquale's remark that "Political dissent is a routine target for surveillance by the FBI", for instance, is not accurate. More often, the book provides appropriately sceptical insights. Did Twitter somehow block the 2011 Occupy Wall Street protests from its own list of high-profile trending topics, critics wondered? The answer is no: trending is a reflection of a relative increase in prominence, not of absolute popularity, as Twitter officials eventually deigned to explain.

The opposite of the black-box society is an "intelligible society", and Pasquale's discussion of it features remedial proposals big and small. Why couldn't the US Library of Congress provide a public book-search function to complement the digitization project Google Books? Why not commission a public credit-scoring system based on open-source software? The underlying question is, why can't the tools of algorithmic decision-making be turned against black-box systems in an open, accountable manner?

It is tempting to say that the political process needed to enact such reforms has itself become a black box. Yet in *The Black Box Society*, Pasquale finds reason to believe that even some of the most secretive and unresponsive institutions can be held to account. Elucidating the problem is a first step. ■

**Steven Aftergood** directs the project on government secrecy at the Federation of American Scientists in Washington DC. e-mail: [saftergood@fas.org](mailto:saftergood@fas.org)



Bill Andrews (left) and Aubrey de Grey have conflicting views on how to delay ageing.

#### AGEING

## Eternal obsession

**Monya Baker** reviews a documentary film profiling two scientists bent on longevity.

Many people want to live forever; few devote their lives to the goal. *The Immortalists*, a documentary film by Jason Sussberg and David Alvarado, tracks two men obsessed with 'curing' ageing. Although their goals are identical, the pair could not be more mismatched. Aubrey de Grey, a geneticist trained at the University of Cambridge, UK, sports a scraggly, chest-length beard and flits around research facilities in Mountain View, California, part-funded by PayPal co-founder Peter Thiel. He regularly pops open a beer at ten in the morning. Bill Andrews, previously a director of molecular biology at biotech company Geron, is a well-trimmed teetotaler who struggles to keep

**The Immortalists**  
DIRECTED BY JASON  
SUSSBERG AND DAVID  
ALVARADO  
Structure Films: 2014.

the lights on at Sierra Sciences, a company he founded in Reno, Nevada.

The first half of the film consists of extended clips of de Grey and Andrews outlining their visions for stalling biological time. Andrews has put his faith in finding small molecules to boost levels of telomerase, the enzyme that restores the fraying ends of chromosomes and declines with age. De Grey has a flamboyant seven-pronged strategy that involves identifying new enzymes and inserting them into a person's body through genetic engineering to create cells that clear out molecular debris. Eventually, the enzymes will be replaced by nanotechnology. Chalkboard-style animated graphics illustrate their key points; the nanobots are particularly cute.

The explanations are framed as straightforward instruction, with no indication of how

MYLEEN HOLLERO

# Correspondence

## Spread excellence across Europe

In February, we expect the results of the first call for proposals for the 'Teaming of Excellence' strategy across the European Research Area. It is to be hoped that we have learned from previous mistakes.

Attempts by the European Union's Framework Programmes for Research and Technological Development to close gaps in research excellence across the region have been spectacularly unsuccessful. In my view, this failure is a result of excessive bureaucracy and a proposal-evaluation system that was seemingly based neither on merit nor on evidence. The misguided focus on upgrading existing research institutions — rather than on creating state-of-the-art structures that are free of the conventional organizational drag on performance — was also a waste of money.

The Horizon 2020 plan for Teaming of Excellence heralds a new approach. It supports the creation of centres of excellence in member states that are lagging behind in research, set up as joint ventures with excellent partners from other member states. It is to be hoped that the commitment to forming these new centres will take priority under the evaluation criteria, so that this time we can genuinely close structural disparities in science and "spread excellence".

**Maciej Żylicz** *Foundation for Polish Science, Warsaw, Poland.*  
[maciej.zylicz@fnp.org.pl](mailto:maciej.zylicz@fnp.org.pl)

## Indirect costs: cash is no gravy train

Aspects of your report on US federal funding of direct research costs and the indirect costs of facilities and administration are misleading (*Nature* **515**, 326–329; 2014).

Contrary to your claim, no one is benefiting from federal largesse. Rather, the

US government is partially reimbursing research universities for audit-verified indirect costs that they have already incurred.

Negotiated and actual recovery rates for indirect costs vary across the academic community because federal research funding is merit-based, not a welfare programme. Moreover, institutions may be private or public, urban or rural, with different structures, sizes, missions and financial anatomies.

The biomedical sciences depend on powerful technologies that require special housing, considerable energy consumption, and maintenance. Administration is being bloated by federal regulations, many of which dictate how scientists conduct and disseminate their research. It is therefore all the more remarkable that the share of extramural research spending on indirect costs by the US National Institutes of Health (NIH) has been stable at around 30% for several decades.

Our respected system of federal funding needs to be accurately and fairly presented in this era of US budget distress and political gridlock, especially given that the NIH budget has lost 25% of its purchasing power in the past decade while grant applications have more than doubled.

**David Korn** *Massachusetts General Hospital and Harvard Medical School, Boston, USA.*  
[david\\_korn@harvard.edu](mailto:david_korn@harvard.edu)

## Indirect costs: the reimbursement gap

The gap between the indirect costs of research reimbursed to US universities and the actual costs is wider than you suggest (*Nature* **515**, 326–329; 2014). If we are to "keep the lights on" to do our research, we must reduce this gap.

For example, the University of South Florida's reimbursed indirect cost rate negotiated from federal funding was 49% of

total direct research expenditure in 2013, whereas the actual rate was more than 54%; the recovery of indirect costs is even lower for many universities after all sources of external research funding are taken into account (K. A. Holbrook and P. R. Sanberg *Technol. Innov.* **15**, 269–280; 2013).

Demands on institutional support for core facilities and for more-flexible funding schemes are likely to increase as university research becomes more transdisciplinary (see [go.nature.com/uwu62q](http://go.nature.com/uwu62q)). Reducing the administrative burden by streamlining processes will help in meeting these demands.

**Paul R. Sanberg, Judy Genshaft, Sudeep Sarkar** *University of South Florida, Tampa, Florida, USA.*  
[psanberg@usf.edu](mailto:psanberg@usf.edu)

## Disaster mapping by citizens is limited

Growing numbers of citizen scientists are joining the professional community to map structural damage caused by natural disasters such as Typhoon Haiyan, which struck the Philippines in 2013 (see *Nature* **515**, 321; 2014). However, stakeholders need to recognize some crucial limitations of crowdsourced contributions.

Citizen scientists are invaluable for mapping roads accurately and for picking out local points of interest. But when it comes to analysing structural damage, even high-resolution satellite data are inadequate. For example, fine-resolution aerial imagery revealed that damage after the 2010 Haiti earthquake was ten times worse than originally concluded by mapping volunteers. Moreover, vertical data cannot provide a comprehensive picture, particularly of partial damage (see N. Kerle and R. R. Hoffman *Nat. Hazards Earth Syst. Sci.* **13**, 97–113; 2013).

Furthermore, volunteers receive only minimal instruction.

They are not trained to map structural damage, which relies on image proxies such as shadows and debris, nor do organizers of the mapping systems provide corrective feedback for improving mapping accuracy.

Coordination with professional organizations on standards and terminology must also be improved if the needs of disaster-response stakeholders are to be met.

**Norman Kerle** *University of Twente, Enschede, the Netherlands.*  
[n.kerle@utwente.nl](mailto:n.kerle@utwente.nl)

## Bold alliances aid translational work

We propose that translational research needs to venture out of its comfort zone and become more interdisciplinary. Physicians, pharmacists, statisticians, computational biologists, social scientists and others from academia, health care and industry should be working hand in hand.

Our interdisciplinary translational research group (see [go.nature.com/sgdrga](http://go.nature.com/sgdrga)), part of an association of research-intensive European institutions in the life sciences called EU-LIFE, aims to foster a continuous cycle from basic to clinical research and back.

To ensure the success of interdisciplinary translational-research programmes, we suggest that each partner should verse the next generation of basic and clinical scientists in the programme's aims; that intellectual-property rights need to be formally stated and managed; that clinicians' research time should be protected; and that funding agencies should create sustainable funding opportunities for basic research that are informed by unmet medical needs.

**Janine T. Erler** *University of Copenhagen, Denmark.*  
[janine.erler@bric.ku.dk](mailto:janine.erler@bric.ku.dk)

*\*On behalf of 4 correspondents (see [go.nature.com/ad21lk](http://go.nature.com/ad21lk) for full list).*



## CATALYSIS

# Gold unleashes the power of three

Gold in the +3 oxidation state is scarcely used in catalysis, because the oxidants employed to generate it can damage reactants. An oxidant-free route to gold(III) catalysts reveals their potential. [SEE ARTICLE P.449](#)

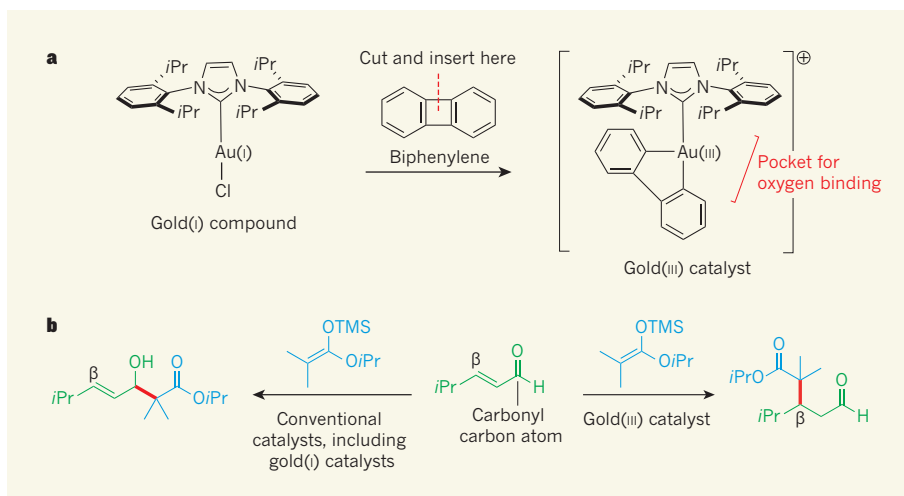
CHRISTOPHER M. B. K. KOURRA &  
NICOLAI CRAMER

A key challenge for organic chemists is to selectively synthesize structurally complex target molecules for the pharmaceutical, agrochemical and materials industries in a rapid, cost-effective manner. Catalysis has attracted considerable attention as a way of achieving this, because of its ability to probe unexplored chemical reactivity, and therefore to enable short synthetic routes to target compounds. On page 449 of this issue, Wu *et al.*<sup>1</sup> report the design and application of catalysts based on gold in the +3 oxidation state (Au(III) catalysts), and describe their use in promoting reactions that are complementary to existing catalytic processes.

Chemists strive to use catalysis to orchestrate the assembly of demanding molecules in a controlled and selective manner. Over the past few decades, game-changing advances in catalysis have focused mainly on transition metals in low oxidation states. By contrast, catalysis that uses transition metals in high oxidation states — such as Au(III) — remains largely underdeveloped<sup>2</sup>.

This situation is particularly evident for gold: Au(I) catalysis<sup>3</sup> is a burgeoning field, behind which Au(III) catalysis lags considerably. The reasons are twofold. First, it is difficult to access the Au(III) oxidation state using ‘mild’ conditions (that is, conditions compatible with the desired catalytic process), and second, Au(III) catalysts have poor stability<sup>4,5</sup>. The successful development of such catalysts therefore requires striking a fine balance between catalytic activity and stability. Progress will continue to be limited until ligand molecules can be developed that stabilize Au(III) metal centres.

Wu and colleagues set out with the lofty ambition of developing a mild strategy that not only allows the generation of stable and active Au(III) catalysts, but also allows the catalysts to be easily prepared and tuned for particular reactions. The central element of their plan was to form Au(III) species by inserting an existing Au(I) catalyst into a carbon–carbon (C–C) bond, a process known as oxidative addition. This avoids the use of any strong oxidants, which are usually required to



**Figure 1 | Preparation and application of a gold(III) catalyst.** **a**, Wu *et al.*<sup>1</sup> report that the gold atom (Au) of a gold(I) compound inserts into a carbon–carbon single bond of biphenylene under mild conditions (at room temperature and without the need for oxidants) to form a gold(III) catalyst. The bulky ligand molecules in the catalyst shield most of the gold atom, but create a pocket in which oxygen atoms of organic molecules can bind to the metal. *iPr*, isopropyl group ( $\text{CH}(\text{CH}_3)_2$ ). **b**, The gold(III) complex promotes catalytic reactivities that are complementary to those obtained using gold(I) catalysts or other conventional catalysts. In this example, a silyl enol ether (blue) reacts with an  $\alpha,\beta$ -unsaturated aldehyde (green) at the carbonyl carbon atom (newly formed bond is shown in red) in the presence of a conventional catalyst, but at the  $\beta$  position in the presence of the gold(III) catalyst. TMS, trimethylsilyl group ( $\text{Si}(\text{CH}_3)_3$ ).

access Au(III) catalysts, and which can be detrimental to chemical groups in the substrates used for the reactions.

The authors knew that several transition metals undergo oxidative addition to a compound called biphenylene<sup>6</sup> — these reactions occur readily because they alleviate strain in biphenylene molecules. Sure enough, when they reacted a Au(I) compound with biphenylene, the metal inserted into the latter's C–C single bond, generating a Au(III) catalyst containing a stabilizing biphenylene ligand (Fig. 1a). The reaction was straightforward to perform and surprisingly mild (occurring at room temperature). The authors used a simple procedure to convert their catalyst into a compound from which the active catalyst itself can be easily generated in reactions. Crucially, this catalyst precursor can be stored and handled easily without the need to exclude moisture and oxygen.

The next task was to apply this Au(III) catalytic system to solve persistent unanswered chemical problems. Au(III) catalysts have a high affinity for oxygen atoms, unlike

their Au(I) counterparts, and thus promote a complementary set of reactions. Wu and co-workers envisaged that they could use the bulky ligands around the gold centre to guide oxygen atoms on reaction substrates into a small, defined pocket in which they would bind to the metal. Crucially, the ligands provide a protective shield that induces the substrate to react with a particular regioselectivity (preferentially at a particular atom), thus avoiding the formation of an undesired mixture of products.

Wu *et al.* tested their catalyst in six types of reaction in which molecules add (form a bond) to the C–C double bonds of compounds known as  $\alpha,\beta$ - and  $\alpha,\beta,\gamma,\delta$ -unsaturated aldehydes. They observed that it overrides the regioselectivity generally obtained using previously reported methods<sup>7,8</sup> (Fig. 1b). In other words, the bulky ligands force substrates to react at only the most remote modifiable position from the aldehyde's oxygen atom. Good product yields were obtained for all the reactions reported, demonstrating the broad applicability and reliability of the Au(III) catalyst.

A particularly exciting element of this work is that it opens up the possibility of using Au(I)- and Au(III)-catalysed processes in tandem, thus combining the best of both worlds. Wu and colleagues exemplified this approach by performing a known reaction<sup>9</sup> using an Au(I) catalyst and then simply adding biphenylene to the reaction mixture, generating the Au(III) catalyst *in situ*. A second reaction catalysed by the Au(III) compound consequently occurred.

The results are a major breakthrough for chemists aiming to master selective reactions at remote positions of molecules that contain several possible sites of modification. The authors' innovative strategy for accessing Au(III) catalysts in a mild, reliable manner, without the use of harsh oxidants, is revolutionary. But there is more to be done. The work currently examines

the reactions of just one class of compound from the large family of molecules that contain carbonyl (C=O) groups. The effects of the Au(III) catalyst on other carbonyl-containing compounds should now be investigated. It also remains to be seen whether the catalyst can be successfully applied to molecules that contain heteroatoms (atoms that are not carbon or hydrogen) other than oxygen. These are exciting times: the door to Au(III) catalysis has been kicked wide open, and chemists are poised to exploit its full potential. ■

**Christopher M. B. Kourra and Nicolai Cramer** are at the *Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Chemical Sciences and Engineering, SB ISIC LCSA, BCH 4305,*

*1015 Lausanne, Switzerland.*  
e-mail: nicolai.cramer@epfl.ch

1. Wu, C.-Y., Horibe, T., Jacobsen, C. B. & Toste, F. D. *Nature* **517**, 449–454 (2015).
2. Hickman, A. J. & Sanford, M. S. *Nature* **484**, 177–185 (2012).
3. Corma, A., Leyva-Pérez, A. & Sabater, M. J. *Chem. Rev.* **111**, 1657–1712 (2011).
4. Gaillard, S., Salwin, A. M. Z., Bonura, A. T., Stevens, E. D. & Nolan, S. P. *Organometallics* **29**, 394–402 (2010).
5. Wolf, W. J., Winston, M. S. & Toste, F. D. *Nature Chem.* **6**, 159–164 (2013).
6. Perthuisot, C. *et al. J. Mol. Catal. A* **189**, 157–168 (2002).
7. Yamamoto, H. (ed.) *Lewis Acids in Organic Synthesis* (Wiley, 2000).
8. Akagawa, K., Sen, J. & Kudo, K. *Angew. Chem. Int. Edn* **52**, 11585–11588 (2013).
9. Pennell, M. N., Unthank, M. G., Turner, P. & Sheppard, T. D. *J. Org. Chem.* **76**, 1479–1482 (2011).

## CELL DIVISION

# Hold on and let go

**The discovery and functional analysis of the protein MEIKIN in mice leads to an evolutionarily conserved model of how chromosome segregation is regulated during a specialized type of cell division called meiosis I. [SEE ARTICLE P.466](#)**

KIKUË TACHIBANA-KONWALSKI

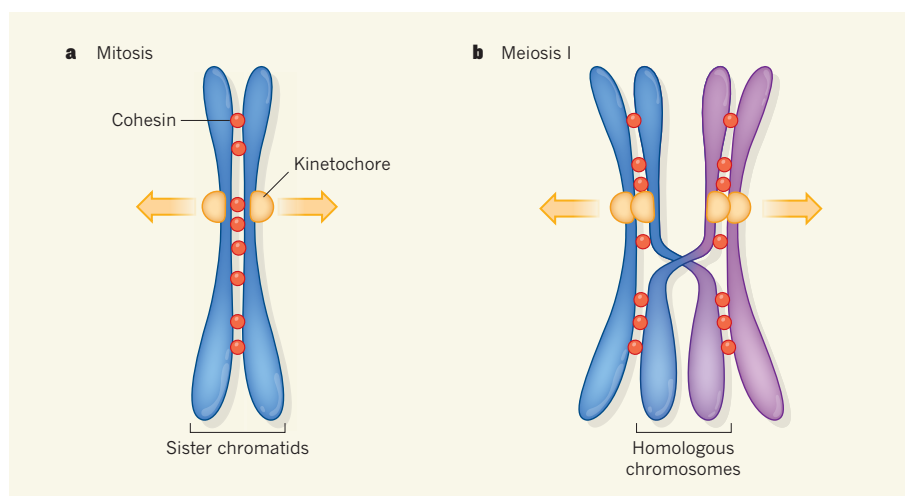
**T**he movements of sister chromosomes during cell division can be compared to those of figure skaters on an ice rink. During most divisions, a pair of skaters stand back-to-back at the centre of the ice and, as the music starts, skate away from each other, separating forever. In a division that leads to the production of eggs or sperm, however, the two skaters stand side-by-side, back-to-back with another couple. Each couple holds hands as the two teams separate, and the skaters make their way across the ice together. Letting go of a partner too early could result in abnormal chromosome segregation and infertility, but the mechanisms that encourage chromosomes to hold on and separate together in this setting have long been a mystery. In this issue, Kim *et al.*<sup>1</sup> (page 466) identify the first key regulator of this process in mammals, illuminating a molecular pathway that seems to be evolutionarily conserved from yeast to humans.

Cell division that produces two identical daughter cells is called mitosis. Before mitosis, each chromosome in the cell is duplicated to produce an identical sister — in humans, 46 chromosomes (23 pairs of 'homologous' chromosomes, one set from each parent) become 92. Duplicated chromosomes called sister chromatids are held together along chromosome arms and at specialized domains, the centromeres, by complexes of cohesin protein. Protein structures called kinetochores are built back-to-back (bi-oriented) on the

centromeres of sister chromatids. Kinetochore bi-orientation facilitates sister-chromatid separation during the anaphase stage of division. At this time, cohesin is destroyed, and spindle fibres that extend from each pole of the cell attach to the facing kinetochore and pull sisters apart, partitioning an identical sister into each daughter cell (Fig. 1a).

By contrast, during meiosis, non-identical sperm or eggs (germ cells) are produced from two rounds of division, meiosis I and II. Meiosis follows chromosome duplication like mitosis, but DNA exchange, known as meiotic recombination, occurs between homologous chromosomes to ensure that germ cells have a mix of genetic material from each parent. Unlike mitosis, cohesin is maintained at centromeres during anaphase of meiosis I. Kinetochore on sister chromosomes face the same direction — they are mono-oriented — and are thus captured and pulled by spindle fibres from the same pole (Fig. 1b). Chromosome segregation in meiosis II proceeds in a similar manner to mitosis, separating the sisters into forming germ cells.

How kinetochore mono-orientation is regulated in meiosis I is a long-standing



**Figure 1 | Chromosome orientation in mitosis and meiosis.** Before cell division, identical chromosome copies called sister chromatids are held together by cohesin protein. How the chromatids segregate during division is determined by the orientation of protein structures called kinetochores — each chromatid is pulled towards the pole of the cell that the kinetochore faces (indicated by arrows). **a**, In mitosis, kinetochores sit back-to-back on sister chromatids (they are bi-oriented), and chromatids are pulled to opposite poles. **b**, In meiosis I, homologous chromosomes are linked as a consequence of meiotic recombination. The kinetochores of the homologous chromosomes face in opposite directions, but those of sister chromatids are mono-oriented and are pulled to the same cell pole.

question. The only proteins known to regulate kinetochore orientation in meiosis specifically are Moa1 in fission yeast and Spo13 and Mam1 in budding yeast. However, the proteins from each species seem to show little genetic similarity to each other<sup>2–5</sup>, making it difficult to identify their mammalian counterparts. Building on the knowledge<sup>6</sup> that Moa1 physically interacts with the kinetochore component protein Cnp3, Kim and colleagues used the equivalent mouse protein, CENP-C, as bait to fish out the corresponding mammalian kinetochore regulator, which they dub MEIKIN (for meiosis-specific kinetochore protein). MEIKIN is at the right place at the right time — located at kinetochores in meiosis I but undetectable in meiosis II.

The authors report that deletion of the *Meikin* gene in mice causes infertility in both sexes, consistent with a defect in meiosis. Tracking of wild-type oocytes (egg-precursor cells) in high spatial and temporal resolution showed that kinetochores of sister chromatids remain in close proximity shortly after anaphase I begins. By contrast, mono-oriented kinetochores split prematurely in MEIKIN-deficient oocytes, and sister chromosomes separate entirely before anaphase of meiosis II.

These results suggest that MEIKIN is directly or indirectly required to maintain cohesin at centromeres. Alternatively, cohesion between centromeres of MEIKIN-deficient oocytes may be only weakly generated in the first place. This is reminiscent of mutants with reduced cohesin gene expression<sup>7</sup>, in which centromere cohesion is abnormal in meiosis II. By inactivating MEIKIN at defined stages of meiosis, it may be possible to delineate whether the protein protects cohesin in anaphase I or is involved in earlier events.

Could MEIKIN regulate kinetochore mono-orientation? If so, why do kinetochores seem to mono-orient in its absence? In yeast, defects in mono-orientation can be obscured when chromosomes are physically linked as a consequence of meiotic recombination<sup>5</sup>. To circumvent this, kinetochore orientation can be analysed in mutant mice in which recombination does not occur. Recombination-deficient oocytes often arrest in meiosis I with mono-oriented kinetochores<sup>8</sup>. The authors demonstrate that, remarkably, loss of MEIKIN in such oocytes can lead to bi-orientation of kinetochores, enabling cell division. Therefore, MEIKIN facilitates kinetochore mono-orientation in addition to protecting centromeric cohesin.

How exactly MEIKIN regulates kinetochore orientation remains to be determined. The protein might help to physically fuse kinetochores together, or it could exert this effect indirectly, regulating as-yet-unidentified 'fusion proteins'. MEIKIN's location on kinetochores suggests that its mode of regulation is likely to be local. It is striking that loss of MEIKIN consistently delays meiosis I by several hours, which could reflect a need

to halt the process until a threshold number of bi-oriented kinetochores is reached and a checkpoint is satisfied. However, it is unclear why this would take so long.

Experiments indicate<sup>9</sup> that the timing defects observed by Kim *et al.* are more severe than those brought about by acute inactivation of centromeric cohesin, which also triggers sister-kinetochore bi-orientation in meiosis I. Therefore, MEIKIN might regulate the cell-cycle machinery. Indeed, Spo13 both promotes kinetochore mono-orientation and regulates APC/C, a protein complex that triggers anaphase<sup>3,4</sup>. Because mouse oocytes are ideal for quantifying the kinetics of APC/C activation, future work may address whether MEIKIN regulates the APC/C, and how this fits with its role at kinetochores.

Both functions ascribed to MEIKIN in this study — maintenance of centromeric cohesion and kinetochore mono-orientation — have previously been attributed to the enzyme PLK1 (refs 10, 11). Kim and colleagues provide evidence that MEIKIN and PLK1 function together. The proteins physically interact, and chemical inhibition of PLK1 has a similar effect to loss of *Meikin*. Furthermore, MEIKIN maintains normal levels of PLK1 at kinetochores. The authors find that this pathway is

conserved in both budding and fission yeast, and that MEIKIN, Moa1 and Spo13 all act through a site that potentially binds PLK1 or its equivalent in yeast. It is this close collaboration that determines whether chromosomes hold on to 'skate' together or separately when the music starts. ■

**Kikuë Tachibana-Konwalski** is at the Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), 1030 Vienna, Austria.  
e-mail: kikue.tachibana@imba.oeaw.ac.at

- Kim, J. *et al.* *Nature* **517**, 466–471 (2015).
- Tóth, A. *et al.* *Cell* **103**, 1155–1168 (2000).
- Lee, B. H., Kiburz, B. M. & Amon, A. *Curr. Biol.* **14**, 2168–2182 (2004).
- Katis, V. L. *et al.* *Curr. Biol.* **14**, 2183–2196 (2004).
- Yokobayashi, S. & Watanabe, Y. *Cell* **123**, 803–817 (2005).
- Tanaka, K., Chang, H. L., Kagami, A. & Watanabe, Y. *Dev. Cell* **17**, 334–343 (2009).
- Murdoch, B. *et al.* *PLoS Genet.* **9**, e1003241 (2013).
- Woods, L. M. *et al.* *J. Cell Biol.* **145**, 1395–1406 (1999).
- Tachibana-Konwalski, K. *et al.* *Curr. Biol.* **23**, 2534–2539 (2013).
- Lee, B. H. & Amon, A. *Science* **300**, 482–486 (2003).
- Katis, V. L. *et al.* *Dev. Cell* **18**, 397–409 (2010).

This article was published online on 24 December 2014.

#### ANTIBIOTICS

## An irresistible newcomer

**A screen of 10,000 bacterial strains, cultured in their normal soil, has uncovered an antibiotic with broad and potent activity. And because the compound targets lipid molecules, developing resistance is probably difficult. SEE ARTICLE P.455**

GERARD WRIGHT

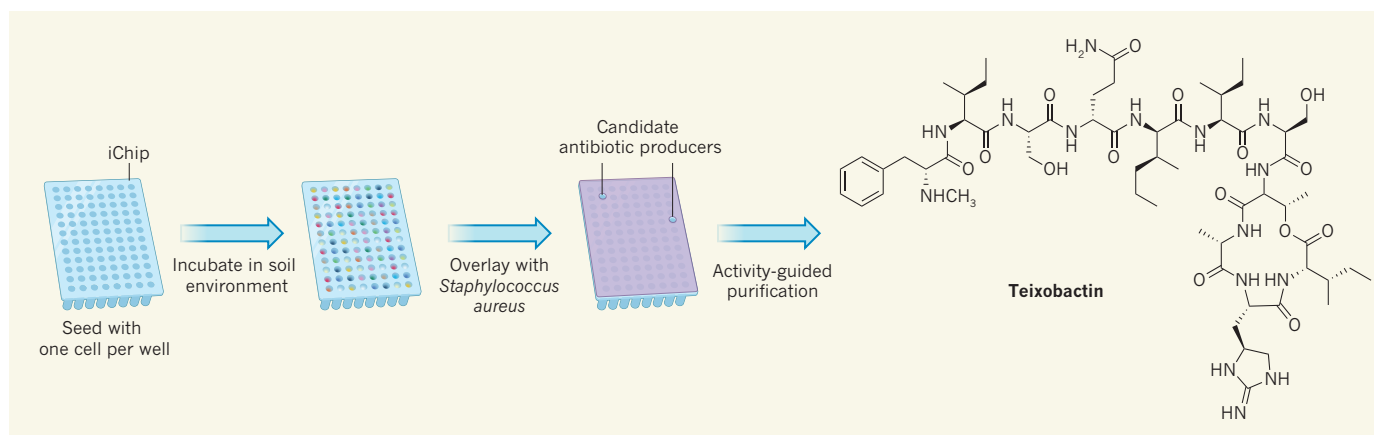
**W**e need new antibiotics. Public-health leaders across the globe are sounding the alarm that the advances of modern medicine are at grave risk in the face of emerging bacterial resistance to existing antibiotic drugs<sup>1,2</sup>. Despite this great clinical need, and widespread recognition of the problem, there are few new antibiotics coming to market, owing to low engagement in this area by the drug-discovery sector and a lack of creative approaches to identifying lead compounds. But on page 455 of this issue, Ling *et al.*<sup>3</sup> present possible solutions to both challenges, with their report of the discovery of an 'uncultivable' soil bacterium that produces an antibiotic that shows no evidence of inducing the evolution of spontaneous resistance in its bacterial targets.

Drug discovery is an expensive business,

and antibiotics, unlike almost all other classes of drug, are inherently vulnerable to the evolution of resistance. Antibiotic discoverers must therefore contend with a built-in 'self-destruct' mechanism that can arise at any time, making return on investment difficult to predict. Furthermore, infectious-disease practitioners and public-health leaders advise against extensive use of new antibiotics, with the aim of slowing the development of resistance, and this results in reduced sales of any new product. In recognition of these challenges, legislative initiatives designed to stimulate commercial investment in antibiotic discovery are being rolled out<sup>4</sup>.

However, alongside this lack of enthusiasm in the pharmaceutical industry is the challenge of selecting starting compounds for the discovery phase of antibiotic development. Most successful antibiotic drugs are microbial natural products or their derivatives and, since the discovery of the first antibiotics in the early





**Figure 1 | Pathway to a new antibiotic.** Ling *et al.*<sup>3</sup> used their previously developed iChip device<sup>10</sup> to cultivate 10,000 soil bacteria in their normal environment. They then overlaid extracts from these bacteria with *Staphylococcus aureus*, to identify extracts containing compounds that could kill this pathogenic bacterium. This identified several

candidate antibiotic-producing bacteria. Using activity-guided purification, in which the candidate extracts were fractionated using several techniques, measuring antibiotic activity along the way, the authors identified the bacterium *Eleftheria terrae* as the producer of a new antibiotic, teixobactin.

twentieth century, soil bacteria and fungi have been our antibiotic wellspring<sup>5</sup>. In fact, most of the chemical scaffolds of the antibiotics in current use were discovered during a period termed the golden age of antibiotics, roughly between 1940 and 1960. But since then, mining these sources for new drugs has been an exercise in diminishing returns, because the same scaffolds are routinely identified in screens. To counter this trend, the pharmaceutical industry turned to synthetic compounds in the 1980s, coupled with target-based screening approaches. But although highly successful in drug discovery for other diseases, this strategy has mostly failed to deliver antibiotics<sup>6,7</sup>, provoking a return to screening microbial natural products<sup>8</sup>.

The problem of such screens repeatedly turning up the same chemical scaffolds may have a lot to do with where we are looking. Actinomycete bacteria, in particular soil-dwelling, spore-forming genera, have long been the mainstay of antibiotic discovery. These are easily grown in the laboratory but represent only a tiny fraction of the microbial genetic diversity in soil; the vast majority of species in environmental samples cannot easily be cultured, and it is thought that this microbial 'dark matter'<sup>9</sup> may represent an untapped source of new antibiotic scaffolds.

One approach to accessing such compounds is through constructing large genomic libraries representing multiple microbial species (metagenomics) and then expressing these genes in other host organisms. Another is the development of new culturing techniques, and the research group presenting the current paper has pioneered several approaches to cultivating rare taxa, including *in situ* growth in natural environments. One of these approaches involves the iChip<sup>10</sup> — a multichannel, miniaturized device for high-throughput cultivation of rare cells directly in their source environments, thereby

enriching for species that require conditions that are not readily replicated using traditional lab methods.

Ling *et al.* used this technology to isolate 10,000 soil bacterial strains, from which they prepared extracts that they examined for the ability to inhibit the growth of *Staphylococcus aureus* — a pathogenic bacterium belonging to the Gram-positive class. One of the extracts showing promising activity was derived from a Gram-negative  $\beta$ -proteobacterium, which the authors name *Eleftheria terrae*. Purification and structural determination of the antibiotic produced by this bacterium, termed teixobactin, revealed a different chemical scaffold from existing antibiotics (Fig. 1). The finding suggests that using such approaches to mine soil microorganisms might uncover a reservoir of new antibiotic scaffolds.

The authors' experiments suggest that teixobactin exerts its antibiotic effects by binding to two precursors of bacterial cell-wall polymers: lipid II (peptidoglycan) and lipid III (teichoic acid). This binding specificity explains the compound's efficacy against Gram-positive bacteria, which have a thick peptidoglycan layer in their teichoic-acid-containing cell wall, and its lack of activity against most Gram-negative bacteria, which are surrounded by an impermeable outer membrane that prevents access to lipid II, and which lack teichoic acids. The authors show that treating mice infected with *S. aureus* or *Streptococcus pneumoniae* (a bacterium that causes upper respiratory tract infections) with teixobactin reduced their infection without obvious toxicity. Furthermore, the researchers' efforts to generate teixobactin-resistant mutants in susceptible Gram-positive bacteria were unsuccessful.

The reason behind the lack of development of teixobactin-resistant bacteria is probably that resistance is less likely to develop against antibiotics that target the lipid molecules essential to cell-wall synthesis than against

antibiotics targeting proteins. Lipids are synthesized by the cell from organic precursors, whereas the genes encoding proteins may mutate, generating structures against which the antibiotic has lower efficacy and giving the mutated cell a survival advantage in the presence of the antibiotic. Vancomycin also binds to lipid cell-wall precursors, and resistance to this antibiotic did not emerge in the clinic until almost 40 years after its discovery in 1953, when the self-resistance mechanism used by vancomycin-producing bacteria was somehow mobilized — probably through horizontal gene transfer — and captured by pathogenic strains<sup>11</sup>. The teixobactin producer identified by Ling *et al.* does not seem to require a specific self-resistance mechanism, owing to the presence of the outer membrane, and so eventual capture of a self-resistance module by other species looks to be unlikely.

It remains to be seen whether other mechanisms of resistance against teixobactin are out there in the environment, but the authors' findings suggest that a systematic search for Gram-negative bacteria producing antibiotics that target the Gram-positive cell wall could identify other 'resistance-light' antibiotics. And reciprocal screens of Gram-positive producers for essential components of the Gram-negative outer membrane might identify similarly promising agents against pathogens of this bacterial class. Thus, in a field dominated by doom and gloom, Ling and colleagues' work offers hope that innovation and creativity can combine to solve the antibiotics crisis. ■

Gerard Wright is in the Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8N 3Z5, Canada.  
e-mail: wrightge@mcmaster.ca

1. Spellberg, B. & Shlaes, D. *Clin. Pharmacol. Ther.* **96**, 151–153 (2014).
2. World Health Organization. *Antimicrobial*

- Resistance: Global Report on Surveillance* (2014).  
 3. Ling, L. L. *et al. Nature* **517**, 455–459 (2015).  
 4. Brown, E. D. *Can. J. Microbiol.* **59**, 153–156 (2013).  
 5. Lewis, K. *Nature Rev. Drug Discov.* **12**, 371–387 (2013).  
 6. Payne, D. J., Gwynn, M. N., Holmes, D. J. &

- Pompliano, D. L. *Nature Rev. Drug Discov.* **6**, 29–40 (2007).  
 7. Silver, L. L. *Clin. Microbiol. Rev.* **24**, 71–109 (2011).  
 8. Wright, G. D. *Can. J. Microbiol.* **60**, 147–154 (2014).  
 9. Rinke, C. *et al. Nature* **499**, 431–437 (2013).  
 10. Nichols, D. *et al. Appl. Environ. Microbiol.* **76**,

- 2445–2450 (2010).  
 11. Marshall, C. G., Lessard, I. A., Park, I. & Wright, G. D. *Antimicrob. Agents Chemother.* **42**, 2215–2220 (1998).

This article was published online on 7 January 2015.

## ASTRONOMY

# Cosmic fog and smog

**It emerges that most of the elements heavier than helium are not found in galaxies, where they can be mixed into future stars and planets. Instead, these elements largely reside far from galaxies in ionized gas and dust particles.**

MOLLY S. PEEPLES

All of the elements on which life is based (carbon, nitrogen, oxygen, iron and so on) are formed in stars and in the explosive stellar deaths known as supernovae. Therefore, it seems reasonable to expect that they will be found where the stars and supernovae are located — in galaxies. However, studies by Shull *et al.*<sup>1</sup> and Peek *et al.*<sup>2</sup> have now revealed that the vast majority of these ‘metals’, as astronomers call all elements not produced in the Big Bang, reside far from the galaxies in which they were born, in the form of both ionized gas and complex molecules.

The realization in 1957 that stars are crucibles for almost all the elements of the periodic table<sup>3</sup> was immediately coupled with the idea

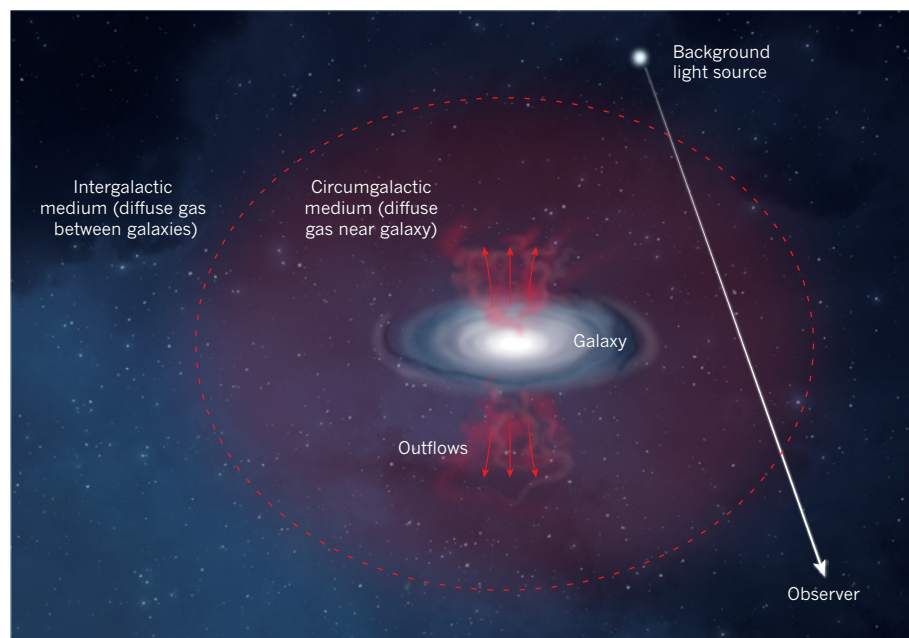
that metals could be used as tracers of star formation and of the flows of gas into, within and out of galaxies (Fig. 1) — much in the way that marked banknotes are used to trace the flows of cash within an economy. Studying metals in galaxies is relatively easy: stars shine brightly, and the gas and dust in galaxies are dense enough to be studied from the light they emit. Yet, as stars age and die, they expel freshly produced material from galaxies, into the circumgalactic medium (the diffuse gas roughly 20 times more extended than the galaxy itself) and perhaps even farther, into the intergalactic medium (the extremely diffuse gas between galaxies). This tenuous gas outside galaxies, however, is so thin that astronomers can observe it only through its effects on the light of bright distant sources passing through

it (Fig. 1). These observations and subsequent interpretations are not easy, because the light passing through the intergalactic medium is affected only minutely by it.

Most of intergalactic space is at densities and temperatures such that the gas consists primarily of ions that most strongly absorb light at ultraviolet (UV) wavelengths. Because Earth’s atmosphere (thankfully) absorbs most UV photons, this means that a space telescope must be used to see this gas in the local, modern-day Universe. Using their outstanding compilation from last year<sup>4</sup> of archival UV spectra of bright background sources observed with the Cosmic Origins Spectrograph aboard the Hubble Space Telescope, Shull *et al.* investigated the carbon, nitrogen, oxygen and silicon littering the pervasive intergalactic medium. These data are mainly sensitive to the intergalactic cosmic densities of metals in certain ionized states (for example, C<sup>2+</sup> and C<sup>3+</sup> but not C<sup>4+</sup>; and O<sup>5+</sup> but not O<sup>6+</sup>).

Because the true intergalactic density of a given element is the sum of its densities in different ionic states, a tricky part of Shull and colleagues’ analysis was correcting these ionic densities for such unseen ionic states to determine to what extent metals have ‘polluted’ intergalactic space. Comparing their newly measured cosmic densities with the expected amounts of metals produced by the stars formed throughout all of cosmic history, Shull *et al.* conclude that 10±5% of all of the metals produced through cosmic time are found in the intergalactic medium, with the rest either still being in galaxies or relatively close by, in the circumgalactic medium.

If Shull and colleagues’ approach to determining the abundances of intergalactic metals using their effects on bright background sources was akin to using a lighthouse to detect cosmic fog, then Peek *et al.* looked at the cosmic smog. They used the Sloan Digital Sky Survey<sup>5</sup> to systematically measure the way in which small intergalactic dust particles (such as graphite, silicate and soot — the same metals that Shull *et al.* observed, but in a solid state) slightly redden background sources by extraordinarily small amounts; this effect is akin to how small particulates in Earth’s atmosphere make the Sun look redder than normal at sunset. By studying how more than 140,000 background galaxies with otherwise similar colours slightly redden depending on the location of foreground galaxies on the sky, Peek *et al.* could beat down statistical uncertainties to measure the profiles of dust in the vicinities of the foreground galaxies. They then used these profiles to calculate the total



**Figure 1 | Observing the diffuse gas outside galaxies.** Stars and supernovae in galaxies are the sites of creation of ‘metals’, but the energy and momentum from dying stars throw these elements (outflows) into the close vicinity of galaxies (the circumgalactic medium) and farther away (into the intergalactic medium). By studying how the light from bright background sources is affected by the diffuse gas outside galaxies, Shull *et al.*<sup>1</sup> and Peek *et al.*<sup>2</sup> have shown that most metals exist far from galaxies, in ionized gas and small, dusty particulates.



- Resistance: Global Report on Surveillance* (2014).  
 3. Ling, L. L. *et al. Nature* **517**, 455–459 (2015).  
 4. Brown, E. D. *Can. J. Microbiol.* **59**, 153–156 (2013).  
 5. Lewis, K. *Nature Rev. Drug Discov.* **12**, 371–387 (2013).  
 6. Payne, D. J., Gwynn, M. N., Holmes, D. J. &

- Pompliano, D. L. *Nature Rev. Drug Discov.* **6**, 29–40 (2007).  
 7. Silver, L. L. *Clin. Microbiol. Rev.* **24**, 71–109 (2011).  
 8. Wright, G. D. *Can. J. Microbiol.* **60**, 147–154 (2014).  
 9. Rinke, C. *et al. Nature* **499**, 431–437 (2013).  
 10. Nichols, D. *et al. Appl. Environ. Microbiol.* **76**,

- 2445–2450 (2010).  
 11. Marshall, C. G., Lessard, I. A., Park, I. & Wright, G. D. *Antimicrob. Agents Chemother.* **42**, 2215–2220 (1998).

This article was published online on 7 January 2015.

## ASTRONOMY

# Cosmic fog and smog

**It emerges that most of the elements heavier than helium are not found in galaxies, where they can be mixed into future stars and planets. Instead, these elements largely reside far from galaxies in ionized gas and dust particles.**

MOLLY S. PEEPLES

All of the elements on which life is based (carbon, nitrogen, oxygen, iron and so on) are formed in stars and in the explosive stellar deaths known as supernovae. Therefore, it seems reasonable to expect that they will be found where the stars and supernovae are located — in galaxies. However, studies by Shull *et al.*<sup>1</sup> and Peek *et al.*<sup>2</sup> have now revealed that the vast majority of these ‘metals’, as astronomers call all elements not produced in the Big Bang, reside far from the galaxies in which they were born, in the form of both ionized gas and complex molecules.

The realization in 1957 that stars are crucibles for almost all the elements of the periodic table<sup>3</sup> was immediately coupled with the idea

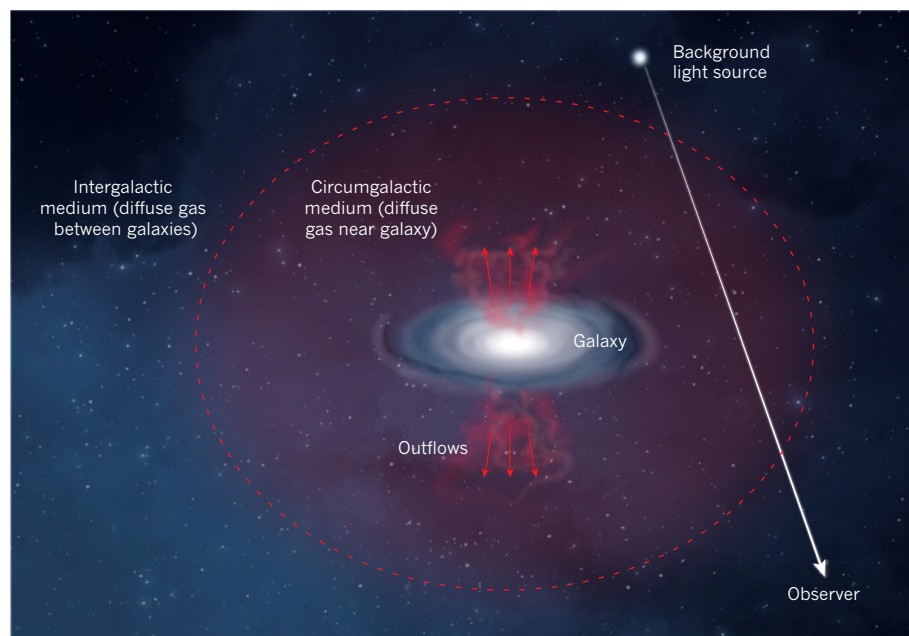
that metals could be used as tracers of star formation and of the flows of gas into, within and out of galaxies (Fig. 1) — much in the way that marked banknotes are used to trace the flows of cash within an economy. Studying metals in galaxies is relatively easy: stars shine brightly, and the gas and dust in galaxies are dense enough to be studied from the light they emit. Yet, as stars age and die, they expel freshly produced material from galaxies, into the circumgalactic medium (the diffuse gas roughly 20 times more extended than the galaxy itself) and perhaps even farther, into the intergalactic medium (the extremely diffuse gas between galaxies). This tenuous gas outside galaxies, however, is so thin that astronomers can observe it only through its effects on the light of bright distant sources passing through

it (Fig. 1). These observations and subsequent interpretations are not easy, because the light passing through the intergalactic medium is affected only minutely by it.

Most of intergalactic space is at densities and temperatures such that the gas consists primarily of ions that most strongly absorb light at ultraviolet (UV) wavelengths. Because Earth’s atmosphere (thankfully) absorbs most UV photons, this means that a space telescope must be used to see this gas in the local, modern-day Universe. Using their outstanding compilation from last year<sup>4</sup> of archival UV spectra of bright background sources observed with the Cosmic Origins Spectrograph aboard the Hubble Space Telescope, Shull *et al.* investigated the carbon, nitrogen, oxygen and silicon littering the pervasive intergalactic medium. These data are mainly sensitive to the intergalactic cosmic densities of metals in certain ionized states (for example, C<sup>2+</sup> and C<sup>3+</sup> but not C<sup>4+</sup>; and O<sup>5+</sup> but not O<sup>6+</sup>).

Because the true intergalactic density of a given element is the sum of its densities in different ionic states, a tricky part of Shull and colleagues’ analysis was correcting these ionic densities for such unseen ionic states to determine to what extent metals have ‘polluted’ intergalactic space. Comparing their newly measured cosmic densities with the expected amounts of metals produced by the stars formed throughout all of cosmic history, Shull *et al.* conclude that 10±5% of all of the metals produced through cosmic time are found in the intergalactic medium, with the rest either still being in galaxies or relatively close by, in the circumgalactic medium.

If Shull and colleagues’ approach to determining the abundances of intergalactic metals using their effects on bright background sources was akin to using a lighthouse to detect cosmic fog, then Peek *et al.* looked at the cosmic smog. They used the Sloan Digital Sky Survey<sup>5</sup> to systematically measure the way in which small intergalactic dust particles (such as graphite, silicate and soot — the same metals that Shull *et al.* observed, but in a solid state) slightly redden background sources by extraordinarily small amounts; this effect is akin to how small particulates in Earth’s atmosphere make the Sun look redder than normal at sunset. By studying how more than 140,000 background galaxies with otherwise similar colours slightly redden depending on the location of foreground galaxies on the sky, Peek *et al.* could beat down statistical uncertainties to measure the profiles of dust in the vicinities of the foreground galaxies. They then used these profiles to calculate the total



**Figure 1 | Observing the diffuse gas outside galaxies.** Stars and supernovae in galaxies are the sites of creation of ‘metals’, but the energy and momentum from dying stars throw these elements (outflows) into the close vicinity of galaxies (the circumgalactic medium) and farther away (into the intergalactic medium). By studying how the light from bright background sources is affected by the diffuse gas outside galaxies, Shull *et al.*<sup>1</sup> and Peek *et al.*<sup>2</sup> have shown that most metals exist far from galaxies, in ionized gas and small, dusty particulates.



mass of dust in the circumgalactic medium of foreground galaxies, concluding that, amazingly, more of these carbon- and silicon-rich molecules reside, not in the galaxies, but outside them.

The combined results of these two studies point to a tantalizing conclusion: most of the elements constituting life are not found in galaxies, where they can be incorporated into future stars and planetary systems; instead, they are predominantly distributed thousands to millions of light years away from galaxies. Although it has been known, or strongly suspected, for decades that galaxies do not contain most of the metals that they have produced, it has been only in recent years that astronomers have been able to systematically locate, quantify and characterize this material outside galaxies.

Both studies, however, highlight how little astronomers know — and thus how much remains to be discovered — about the detailed

physical conditions and hence the fates of the Universe's metals. Shull and co-workers have produced an estimate of the intergalactic density of metals that is for a time when the Universe is much more mature than the earlier epochs at which similar previous measurements<sup>6,7</sup> were made. However, it is not a trivial task to connect these newly detected metals to the individual galaxies in which they were born. Although previous studies<sup>8</sup> have indicated that there is a significant reservoir of intergalactic dust, Peek and colleagues' analysis exquisitely shows how a galaxy's dusty environment depends on the galaxy itself. Large surveys will be required to obtain similar measurements at earlier cosmic epochs, to decipher when galaxies expelled this dust.

It is astounding that most of the potential building blocks of life are found so far from their birthplaces, and that so much of it can survive the trip out of galaxies as complex molecules. As our understanding of intergalactic

space in the local Universe increases, so does what this vast expanse of rarefied gas tells us about galaxies — and about just how rare and precious our own place in the cosmos is. ■

**Molly S. Peebles** is at the *Space Telescope Science Institute, Baltimore, Maryland 21218, USA.*  
e-mail: [molly@stsci.edu](mailto:molly@stsci.edu)

1. Shull, J. M. *et al. Astrophys. J.* **796**, 49 (2014).
2. Peek, J. E. G., Ménard, B. & Corrales, L. Preprint at <http://arxiv.org/abs/1411.3333> (2014).
3. Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. *Rev. Mod. Phys.* **29**, 547–650 (1957).
4. Danforth, C. W. *et al.* Preprint at <http://arxiv.org/abs/1402.2655> (2014).
5. Abazajian, K. N. *et al. Astrophys. J. Suppl.* **182**, 543–558 (2009).
6. Cooksey, K. L. *et al. Astrophys. J.* **708**, 868–878 (2010).
7. Cooksey, K. L. *et al. Astrophys. J.* **762**, 37–52 (2013).
8. Ménard, B. *et al. Mon. Not. R. Astron. Soc.* **405**, 1205–1039 (2010).

## PALAEOCLIMATE

# Monsoon matters

**A simplified global climate model that keeps track of water as it moves through Earth's water cycle throws fresh light on how the Asian summer monsoon has varied during the past 150,000 years.**

BRONWEN KONECKY

The Asian summer monsoon (ASM) is the predominant source of rainy-season precipitation in tropical Asia, and is known to vary in response to global climate changes<sup>1</sup>. However, different geological archives from South and East Asia, such as cave deposits and marine sediments, provide conflicting accounts of how the ASM responds to basic climate forcings (Fig. 1) — for example, varying concentrations of greenhouse gases, the growth and collapse of ice sheets and changes in Earth's orbit around the Sun. Writing in *Nature Communications*, Caley *et al.*<sup>2</sup> use a simplified model to simulate Earth's climate during the past 150,000 years, and find that this conflict may originate from our interpretation of climate signals recorded in stalagmites from Chinese caves.

On timescales of thousands to millions of years, changes in the tilt of Earth relative to the plane of its orbit, the shape of the orbit and its slow wobble about its axis (or 'precession') regulate and distribute incoming solar radiation (insolation) that fuels the ASM. The monsoon is thought to intensify during minima in precession, approximately every 23,000 years, when Earth is closest to the Sun during Northern Hemisphere summer (unlike today, when this occurs in winter). Further strengthening

of the ASM occurs in part as a result of the growth and collapse of ice sheets covering northern Eurasia, which alter key wind patterns and shift the amount and location of East Asian rainfall. Greenhouse gases also influence the ASM by increasing the build-up



**Figure 1 | The Indo-Asian monsoon system.** The shaded branched arrow denotes prevailing monsoonal wind flow. Red dots represent locations of cave stalagmites, which provide records of past climate. Caley *et al.*<sup>2</sup> question whether these stalagmites truly reflect variations in the Asian summer monsoon, in contrast to previous interpretations.

of moisture in the atmosphere that is available to feed monsoonal precipitation.

But how can we tell what caused ASM variations in the geological past? When decoding Earth's climate history, timing is everything. The timing of semi-regular changes in rainfall relative to that of insolation, ice sheets and greenhouse-gas concentrations tells us which mechanisms played what part, and when. In many geological archives, ASM maxima seem to result from a combination of three factors: orbitally driven heating of the Asian land-mass; release of energy in the southern Indian Ocean, where monsoon moisture originates; and minima in global ice volume. By contrast, stalagmite archives from South and East Asian caves — arguably the best-dated, highest-resolution, most continuous records of Asian palaeoclimate recovered so far — suggest that local insolation is the dominant factor, with changes in global ice volume being secondary. This apparent contradiction in timing has long challenged scientists researching the underlying controls on the ASM.

Like many rainfall proxies from the tropics, cave stalagmite records are based on the ratios of stable oxygen and hydrogen isotopes in water. The abundances of heavy but rare <sup>18</sup>O and <sup>2</sup>H relative to their lighter, more common counterparts (expressed as  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$ , respectively) carry an imprint of that water's history as it evaporated from the ocean, moved through the water cycle and eventually fell as rain or snow. When water gets incorporated into the calcite of cave stalagmites, it carries with it the isotopic signature that it had when it fell as precipitation.

Precipitation  $\delta^{18}\text{O}$  is a proven tracer of hydrological and atmospheric circulation processes<sup>3</sup>. However, studies<sup>4,5</sup> have shown that interpreting precipitation  $\delta^{18}\text{O}$  is not always straightforward, because it integrates many

processes across space and time. 'Isotope-enabled' climate models<sup>6</sup> are quickly gaining prominence as tools for deciphering climate signals contained in  $\delta^{18}\text{O}$ . But even with state-of-the-art supercomputers, running a climate model over many thousands of years is expensive, and adding in the extra code to trace water isotopes only steepens that cost.

Caley *et al.* solve this dilemma by using an 'intermediate-complexity' climate model and by accelerating climate forcings by a factor of ten to reduce the amount of time needed to run the simulation. The model, called *iLOVECLIM*, involves a simple, three-layer atmosphere, together with basic isotope physics to describe the fractionation (separation) of O and H isotopes during precipitation and evaporation. These isotopic signatures are then transported through the atmosphere, ocean and land. The model's coarse resolution blurs Earth's topography and neglects complex dynamics of clouds and convection, but its simplicity makes it computationally affordable, enabling the authors to perform multiple 150,000-year experiments.

Despite its simplicity, *iLOVECLIM* reproduces long-term variations in Chinese-cave  $\delta^{18}\text{O}$  very well, albeit with a few telling exceptions. Although Caley *et al.* find that precessional and ice-sheet forcing can explain most cave  $\delta^{18}\text{O}$  variations, *iLOVECLIM* does not quite capture the observed timing offsets between precipitation, cave  $\delta^{18}\text{O}$  and precession minima. This suggests that relevant climate forcings or feedbacks are missing, including 'known unknowns' such as ice-ocean feedbacks that were deliberately left out of the model, but possibly other processes as well.

Importantly, the authors argue that annual  $\delta^{18}\text{O}$  variations in *iLOVECLIM* match Chinese-cave records better than summer  $\delta^{18}\text{O}$  variations — contradicting the common interpretation of cave records as an ASM proxy. Modern data support the idea that annual rainfall  $\delta^{18}\text{O}$  in China reflects both monsoonal and non-monsoonal processes<sup>7</sup>. Non-monsoonal sources of moisture, along with their transport paths and the types of precipitation they produce, account for a large proportion of  $\delta^{18}\text{O}$  variability<sup>5,8</sup>. Infiltration of water through limestone en route to a cave further integrates rainwater  $\delta^{18}\text{O}$  signals over many seasons<sup>9</sup>. Hence, ASM rainfall drives much of the year-to-year variability in annual  $\delta^{18}\text{O}$ , but it is not the only contributor.

The complexity of precipitation  $\delta^{18}\text{O}$  highlights the need to model precipitation and isotopic fractionation in monsoonal Asia realistically. Such modelling requires more-sophisticated atmospheric physics and dynamics than a simplified model can provide. *iLOVECLIM* demonstrates this point by poorly simulating young (thousands of years old) cave  $\delta^{18}\text{O}$  deposits in South and East Asia, mainly because the model oversimplifies precipitation  $\delta^{18}\text{O}$ . Hence, offsets between

Chinese-cave  $\delta^{18}\text{O}$  and modelled-cave  $\delta^{18}\text{O}$  — both summer and annual — may arise from real-world, season-specific precipitation and isotopic phenomena that *iLOVECLIM* does not capture. The proportion of long-term  $\delta^{18}\text{O}$  variability that can be attributed to the ASM, then, remains difficult to determine.

By explicitly modelling water isotopes, Caley *et al.* have taken a crucial first step towards understanding how global climate forcings drive the ASM. Further deciphering  $\delta^{18}\text{O}$  and resolving apparent contrasts in proxy data will require a systems approach: modelling the climate signals themselves, as well as the archives that preserve them<sup>10</sup>. Once the controls on precipitation  $\delta^{18}\text{O}$  have been identified, the underlying causes of ASM variations will come into focus. ■

Bronwen Konecky is at the Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309, USA, and at the College of

Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis.  
e-mail: bronwen.konecky@colorado.edu

1. Christensen, J. H. *et al.* in *Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) Ch. 14, 1217–1308 (Cambridge Univ. Press, 2013).
2. Caley, T., Roche, D. M. & Renssen, H. *Nature Commun.* **5**, 5731; <http://dx.doi.org/10.1038/ncomms6371> (2014).
3. Dansgaard, W. *Tellus* **16**, 436–468 (1964).
4. Aggarwal, P. K., Frölich, K., Kulkarni, K. M. & Gourcy, L. L. *Geophys. Res. Lett.* **31**, L08203 (2004).
5. Dayem, K. E., Molnar, P., Battisti, D. S. & Roe, G. H. *Earth Planet. Sci. Lett.* **295**, 219–230 (2010).
6. Noone, D. & Sturm, C. in *Isoscapes: Understanding Movement, Pattern, and Process on Earth Through Isotope Mapping* (eds West, J. B., Bowen, G. J., Dawson, T. E. & Tu, K. P.) 195–219 (Springer, 2010).
7. Clemens, S. C., Prell, W. L. & Sun, Y. *Paleoceanography* **25**, PA4207 (2010).
8. Lee, J.-E. & Fung, I. *Hydrol. Process.* **22**, 1–8 (2008).
9. Moerman, J. W. *et al.* *Geophys. Res. Lett.* **41**, 7907–7915 (2014).
10. Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M. & Anchukaitis, K. J. *Quat. Sci. Rev.* **76**, 16–28 (2013).

## CELL BIOLOGY

## On the endocytosis rollercoaster

**Endocytosis is a process by which molecules gain access to a cell. An unusual mode of endocytosis has now been shown to regulate cell signalling, and to be hijacked by bacterial toxins. SEE ARTICLE P.460 & LETTER P.493**

VOLKER HAUCKE

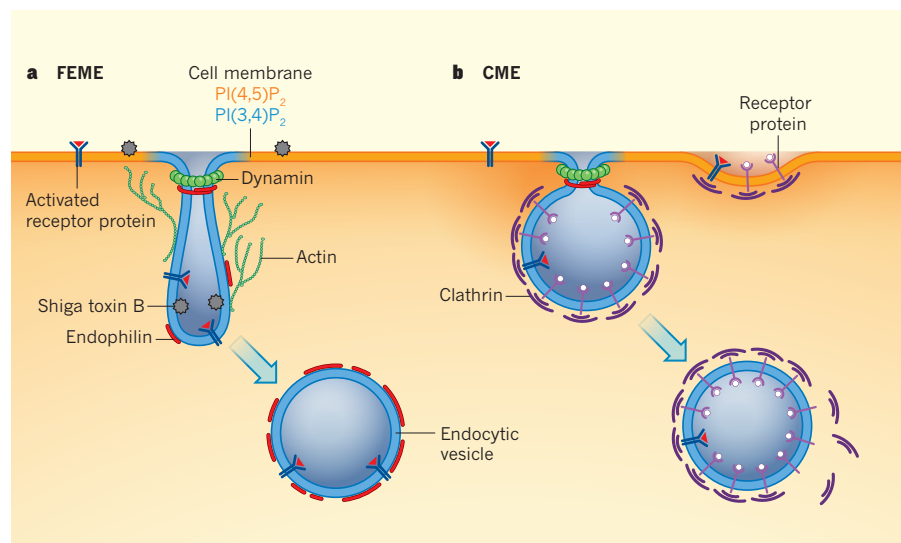
The plasma membrane that surrounds cells forms a barrier that gates access to the cell, permitting entry to nutrients and extracellular messenger molecules, but locking out hazardous compounds and deadly viruses<sup>1</sup>. Clathrin protein coats some regions of this membrane, and controls cellular entry of beneficial molecules through a process called clathrin-mediated endocytosis<sup>1</sup>. Cells are also thought to use clathrin-independent modes of endocytosis<sup>2</sup>, but these have proved difficult to pinpoint. In this issue, Boucrot *et al.*<sup>3</sup> (page 460) describe a fast, clathrin-independent pathway for the import of activated receptor proteins. In a separate study, Renard *et al.*<sup>4</sup> (page 493) show that this pathway is hijacked by toxins from bacteria of the genera *Shigella* or *Vibrio* (which causes cholera), to allow them to gain entry to the cell.

During clathrin-mediated endocytosis (CME), nutrients, hormones, or other ligands that bind to receptor proteins on clathrin-coated membrane regions, are shuttled into cells in small vesicles, which form from invaginations in the membrane. CME not only

controls uptake of receptor-bound molecules, but also serves a general function in regulating the turnover of many membrane-bound proteins. In addition to CME, the cell uses several atypical, clathrin-independent routes of endocytosis. Molecules thought to be imported through these pathways include bacterial toxins such as Shiga toxin B (ref. 5), from *Shigella*, and certain membrane-bound receptor proteins<sup>2,6</sup> that, when on the cell surface, mediate transmission of signals from extracellular signalling factors to the cell nucleus.

Boucrot and colleagues set out to investigate the molecular nature of clathrin-independent endocytosis by studying a membrane-deforming protein, endophilin. Previous genetic analysis of mice lacking all three isoforms<sup>7</sup> of endophilin shows that one function of this protein is to promote clathrin-coat shedding from vesicles during the late stages of CME. However, there is evidence that endophilin can also interact with membrane-bound receptors — either directly, as with the  $\beta_1$ -adrenergic receptor<sup>8</sup>, or indirectly through an intermediate adaptor protein, as with certain growth-factor receptors<sup>9,10</sup>.

The authors confirmed that endophilin



**Figure 1 | Mechanisms of endocytosis.** **a**, Boucrot *et al.*<sup>3</sup> describe a method of cell entry dubbed fast endophilin-mediated endocytosis (FEME), which is triggered by activation of membrane-bound receptor proteins, or by bacterial toxins such as Shiga toxin B. The pathway begins with the conversion of the phospholipid molecule PI(4,5)P<sub>2</sub> to PI(3,4)P<sub>2</sub>, which leads to recruitment of the protein endophilin to the cell membrane. The membrane invaginates into tubules, which are subsequently snipped to form vesicles. Renard *et al.*<sup>4</sup> show that this process requires the coordinated activity of endophilin, dynamin and actin proteins. **b**, By contrast, in clathrin-mediated endocytosis (CME), receptors (including the activated receptors that are involved in FEME) and other cargo are captured into clathrin-protein-coated membrane invaginations, which then undergo dynamin-mediated scission. Clathrin is subsequently removed from the vesicle surface. Like FEME, CME requires PI(3,4)P<sub>2</sub>, but this is produced through a different mechanism<sup>10</sup>.

can directly or indirectly associate with many activated receptors. For example, they found that activation of the  $\beta_1$ -adrenergic receptor triggers the recruitment of endophilin to clathrin-free membrane sites at the cell's leading edge (the edge at the front of the cell during migration). This leads to the rapid internalization of the receptor into small vesicles and tubules. The researchers named this entry mechanism fast endophilin-mediated endocytosis (FEME), owing to its apparent speed.

FEME could be abrogated by depleting endophilin, but not by depleting or inhibiting clathrin and its partner proteins, indicating that this mode of entry is not clathrin-dependent (Fig. 1). Furthermore, the vesicles formed during FEME seem to be distinct from those formed during other clathrin-independent modes of endocytosis, such as macropinocytosis. In the absence of endophilin, receptors accumulate on the cell surface and continue to signal, instead of being internalized — this finding may be important, because increased growth-factor signalling through such receptors is a hallmark of cancer.

Impressively, Boucrot and colleagues determined the mechanism by which receptor activation causes endocytosis. First, activation triggers the addition of a phosphate group to phosphatidylinositol 4,5-bisphosphate (PI(4,5)P<sub>2</sub>) — a membrane-bound phospholipid molecule that activates many cell-signalling pathways — to form another phospholipid, PI(3,4,5)P<sub>3</sub>. This transfor-

mation is transient, because PI(3,4,5)P<sub>3</sub> is rapidly converted to PI(3,4)P<sub>2</sub>, which binds to a protein called lamellipodin that is distributed at the cell's leading edge. Lamellipodin binds endophilin, and thus endophilin accumulates at the leading edge. Accumulated endophilin induces the formation of vesicles that then become internalized, transporting their cargo to the cell's centre.

Renard *et al.*<sup>4</sup> convincingly demonstrate that Shiga and cholera toxins hijack the FEME pathway. They show that the A2 isoform of endophilin coats tubule-shaped membrane invaginations that arise when Shiga toxin B binds to the cell membrane. A similar endophilin-based coat may also form in receptor-activated FEME, although this was not investigated. The authors further show that, during FEME, both endophilin and dynamin (another protein involved in endocytosis) cooperate with actin, a structural protein that can polymerize into branched filaments, to snip off the tubule from the membrane, internalizing Shiga toxin B. Such severing probably occurs through a mechanism based on friction, which may be induced by interactions between the endophilin coat and the underlying lipids as the tubule extends and actin polymerizes.

For the first time, the molecular machinery that controls a specific clathrin-independent mode of endocytosis has been defined. FEME might be required in many cell types, including neurons, which use both CME and clathrin-independent endocytosis to recycle the vesicles

responsible for the release of neurotransmitter molecules at neuronal junctions (synapses)<sup>11</sup>. A notable feature of FEME is that it must be triggered to occur, either by receptor-activated signalling cascades or by toxin-induced membrane deformation. By contrast, CME is believed to be largely constitutive.

FEME is unlikely to be the only pathway of clathrin-independent endocytosis, as Boucrot *et al.* note. They show that, in the absence of either CME or FEME, growth-factor receptors such as epidermal growth factor receptor can still be internalized when activated, through macropinocytosis. It will also be interesting to determine the relative contributions of FEME and CME to general membrane recycling under physiological conditions, because previous studies have yielded conflicting results regarding the relative contributions of CME<sup>12</sup> and clathrin-independent pathways<sup>2</sup> to this process.

The current findings open up avenues for further mechanistic studies. It is unclear how endophilin is recruited to either the CME or the FEME pathway, for instance in synapses, where both processes occur. Lamellipodin might control recruitment for FEME, but the protein's equivalent in CME is unknown. Furthermore, FEME and CME share several components, including dynamin and PI(3,4)P<sub>2</sub> (although this lipid is synthesized by distinct routes in FEME and CME<sup>13</sup>). This not only presents possibilities for simultaneous pharmacological inhibition of both pathways, but also indicates that cells can reuse functional protein modules, for example the proteins involved in membrane scission, and can direct them to more than one pathway. Such direction might depend on environmental conditions, for example cell type or physiological stimulus. Future studies will be required to fully understand the rollercoaster ride of endophilin-mediated endocytosis. ■

**Volker Haucke** is at the Leibniz Institut für Molekulare Pharmakologie, 13125 Berlin, Germany.  
e-mail: haucke@fmp-berlin.de

1. Doherty, G. J. & McMahon, H. T. *Annu. Rev. Biochem.* **78**, 857–902 (2009).
2. Howes, M. T., Mayor, S. & Parton, R. G. *Curr. Opin. Cell Biol.* **22**, 519–527 (2010).
3. Boucrot, E. *et al. Nature* **517**, 460–465 (2015).
4. Renard, H.-F. *et al. Nature* **517**, 493–496 (2015).
5. Römer, W. *et al. Cell* **140**, 540–553 (2010).
6. Lamaze, C. *et al. Mol. Cell* **7**, 661–671 (2001).
7. Milosevic, I. *et al. Neuron* **72**, 587–601 (2011).
8. Tang, Y. *et al. Proc. Natl Acad. Sci. USA* **96**, 12559–12564 (1999).
9. Soubeyran, P., Kowanez, K., Szymkiewicz, I., Langdon, W. Y. & Dikic, I. *Nature* **416**, 183–187 (2002).
10. Petrelli, A. *et al. Nature* **416**, 187–190 (2002).
11. Kononenko, N. L. *et al. Neuron* **82**, 981–988 (2014).
12. Bitsikas, V., Correa, I. R. & Nichols, B. J. *eLife* **3**, e03970 (2014); <http://dx.doi.org/10.7554/eLife.03970>
13. Posor, Y. *et al. Nature* **499**, 233–237 (2013).

This article was published online on 17 December 2014.



# Stable gold(III) catalysts by oxidative addition of a carbon–carbon bond

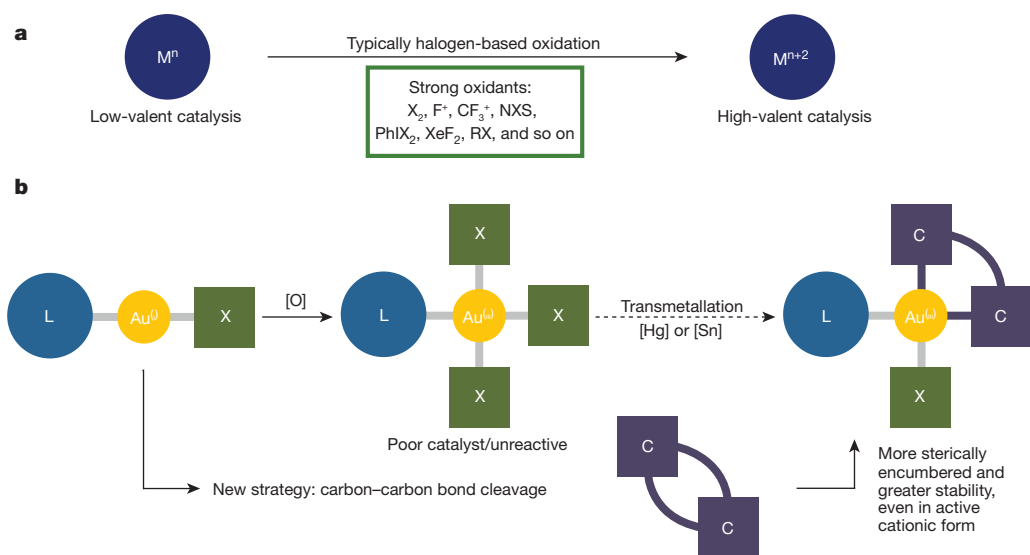
Chung-Yeh Wu<sup>1\*</sup>, Takahiro Horibe<sup>1\*</sup>, Christian Borch Jacobsen<sup>1</sup> & F. Dean Toste<sup>1</sup>

Low-valent late transition-metal catalysis has become indispensable to chemical synthesis, but homogeneous high-valent transition-metal catalysis is underdeveloped, mainly owing to the reactivity of high-valent transition-metal complexes and the challenges associated with synthesizing them. Here we report a carbon–carbon bond cleavage at ambient conditions by a Au(I) complex that generates a stable Au(III) cationic complex. In contrast to the well-established soft and carbophilic Au(I) catalyst, this Au(III) complex exhibits hard, oxophilic Lewis acidity. For example, we observed catalytic activation of  $\alpha,\beta$ -unsaturated aldehydes towards selective conjugate additions as well as activation of an unsaturated aldehyde–allene for a [2 + 2] cycloaddition reaction. The origin of the regioselectivity and catalytic activity was elucidated by X-ray crystallographic analysis of an isolated Au(III)-activated cinnamaldehyde intermediate. The concepts revealed suggest a strategy for accessing high-valent transition-metal catalysis from readily available precursors.

Transition-metal catalysis has developed into an efficient and selective strategy for organic transformations in modern chemistry. Low-valent late (that is, their periodic table group is more to the right) transition-metal complexes are particularly heavily used owing to their stability and usefulness in forming important chemical bonds (C–C, C–O, C–N). However, low-oxidation-state, late transition metals are less suitable for other critical reactions, including electrophilic C–H functionalization<sup>1,2</sup>. Recent efforts have begun to unlock the potential of high-valent late transition metals, especially Pd(IV), to complement these shortcomings<sup>3–6</sup>. The major challenges thus far include the typical need for strong oxidants to access the higher oxidation state, which limits the functional group tolerance, and the instability of the oxidized metal complexes, which often exist only as high-energy intermediates on the catalytic cycle

(Fig. 1a). Put broadly, to fully explore the undoubtedly rich chemistry of high-valent late transition metals, there is a need to develop easily prepared, stable, tunable catalysts.

In this regard we searched for a route to stable, catalytically active Au(III) complexes with the goal of complementing the ever-increasing library of Au(I)-catalysed reactions. Whereas homogeneous Au(I) catalysis has seen great progress over the last decade<sup>7–11</sup>, Au(III) catalysis is still mainly limited to the use of inorganic Au(III) salts<sup>12</sup>. The synthetic challenge in forming stable, yet catalytically active, organometallic Au(III) complexes derives from the intrinsically high redox potential, leading to the facile reduction of Au(III) complexes to Au(I) or metallic Au(0) species in the presence of electron-rich reagents<sup>13–18</sup>. In cases where the ligands are capable of stabilizing the highly oxidizing metal, the resulting complex is often rendered



**Figure 1 | Routes to high-valent metal complexes.** **a**, Standard route to high-valent metal complexes using halogen-based oxidants.  $n$ , oxidation state;  $NXS$ , N-halosuccinimide;  $M$ , metal. **b**, Hypothetically synthetic pathway for accessing Au(III)-C bond-stabilized Au(III) complexes.  $X$ , heteroatom;  $L$ , ligand.

<sup>1</sup>Department of Chemistry, University of California, Berkeley, California 94720, USA.

\*These authors contributed equally to this work.

catalytically inert. For example, the oxidative addition product,  $L\text{-AuX}_3$ , formed from  $L\text{-AuX}$  (where  $X$  is a halogen and  $L$  is ligand) is a poor catalyst by itself, and the abstraction of a halide to increase reactivity enables a facile reduction to the lower oxidation states<sup>15,16</sup>. Moreover, although complexes of the type  $\text{Au(III)}(\text{C}^\wedge\text{L})(\text{X})(\text{Y})$  ( $\text{C}^\wedge\text{L}$  is a bidentate ligand coordinating one carbon and one two-electron charge-neutral ligand;  $X$ , halogen;  $Y$ ,  $X$  or  $L$ ) are available through multi-step synthetic sequences<sup>19–21</sup>, their instability in the cationic form and the difficulties in readily tuning the coordination environment has severely limited their applications in catalysis.

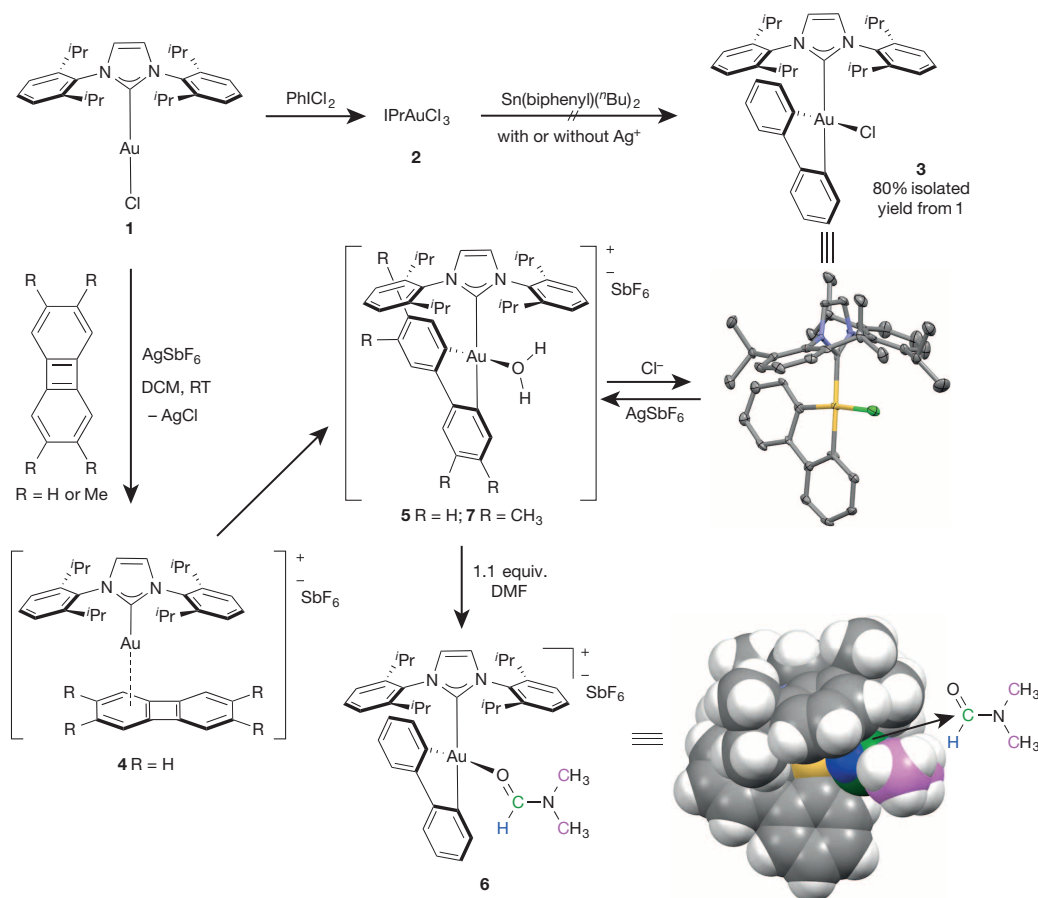
### Formation of stable gold(III) complexes

To address this challenge, we hypothesized that a multidentate ligand scaffold with strong  $\text{Au-C}$  bonding energy might be able to stabilize cationic  $\text{Au(III)}$  organometallic complexes while maintaining catalytic activity. In designing the desired complex, we also hoped to avoid using strong oxidants to maximize the functional group compatibility of the protocol.

From reported examples<sup>22</sup>, we reasoned that the transmetalation of  $\text{Sn}(\text{biphenyl})(^t\text{Bu})_2$  (where  $\text{Bu}$  is butyl) to  $L\text{-AuX}_3$  might afford a stable  $\text{Au(III)}$  complex with a bidentate ligand containing two strong  $\text{Au-C}$  bonds (Fig. 1b). The known stabilization of both neutral and cationic  $\text{Au(I)}$  complexes by  $N$ -heterocyclic carbenes inspired us to examine them as supporting ligands<sup>23</sup>. In an effort to access such compounds, attempts to perform the transmetalation of  $\text{Sn}(\text{biphenyl})(^t\text{Bu})_2$  to  $\text{IPr-AuCl}_3$  (compound **2**) (where  $\text{IPr}$  is [1,3-bis(2,6-diisopropylphenyl)]

imidazol-2-ylidene]) gave no desired product (Fig. 2). Although the replacement of  $\text{IPrAuCl}_3$  with (tetrahydrothiophene) $\text{AuCl}_3$  enabled access to complex **3** in good yield, the two-step procedure was cumbersome, and did not meet the goals of being mild and straightforward.

As an alternative, we imagined that insertion of an  $\text{Au(I)}$  complex into the strained  $\text{C-C}$  bond of biphenylene could achieve the desired oxidation and introduction of a stabilizing biphenyl ligand in a single step<sup>24–26</sup> (Fig. 1b). The concept was appealingly simple, even though oxidative addition to  $\text{Au(I)}$  complexes with mild oxidants is exceedingly rare, and in general requires either bimetallic complexes generating  $\text{Au(II)-Au(II)}$  intermediates (for examples see ref. 27 and references therein, and refs 18 and 28) or complexes with specially designed ligands<sup>29,30</sup>. Moreover, no well defined carbon-carbon bond cleavage reactions with  $\text{Au(I)}$  complexes have been reported. To investigate the feasibility of this strategy, the coordinatively unsaturated  $\text{IPrAu(I)}(\text{SbF}_6)$  was generated *in situ* by reacting  $\text{IPrAuCl}$  with  $\text{AgSbF}_6$ , resulting in the precipitation of  $\text{AgCl}$  in  $\text{CD}_2\text{Cl}_2$  at room temperature. Reaction of the resulting electrophilic  $\text{IPrAu(I)}(\text{SbF}_6)$  species with biphenylene resulted in the immediate and quantitative formation of the coordination complex ( $\text{IPrAu-biphenylene}$ ) ( $\text{SbF}_6$ ) (compound **4** in Fig. 2), along with a small amount of the  $\text{Au(III)}$  aqua complex  $[\text{IPrAu(III)}(\text{biphenyl})(\text{H}_2\text{O})]\text{SbF}_6$  (compound **5**), both of which were observed by  $^1\text{H}$  nuclear magnetic resonance (NMR) spectroscopy<sup>31</sup>. Nearly full conversion of the  $\text{Au(I)}$  cationic species to the desired  $\text{Au(III)}$  complex **5** was observed after 1.5 h (see Supplementary Information, Supplementary Fig. 1 and Supplementary Table 1). In contrast, most reported examples of this type of  $\text{C-C}$  bond cleavage requires more



**Figure 2 | Accessing  $\text{Au(III)}$  via oxidative addition of a carbon-carbon bond.** Attempted access to complex **3** via transmetalation of  $\text{Sn}(\text{biphenyl})(^t\text{Bu})_2$  with  $\text{IPrAuCl}_3$  (compound **2**) is shown. ( $\text{DCM}$ , dichloromethane;  $\text{RT}$ , room temperature (approx.  $23^\circ\text{C}$ );  $-\text{AgCl}$ , removal of  $\text{AgCl}$ ). The proposed pathway for the oxidative addition of  $\text{IPrAu(I)}$  complex with biphenylene/2,3,6,7-tetramethylbiphenylene is shown. The coordination chemistry of  $\text{IPrAu(III)}(\text{biphenyl})$  complexes **3** and **6** is shown. The X-ray

structures of  $\text{IPrAu(III)}(\text{biphenyl})\text{Cl}$  (complex **3**) and  $[\text{IPrAu(III)}(\text{biphenyl})(\text{Me}_2\text{NC}(\text{O})\text{H})][\text{SbF}_6]$  (complex **6**) are shown; the  $\text{SbF}_6^-$  anion is omitted for clarity. In the crystal structure of **3**: gold, Au; green, Cl; blue, N; grey, C; H is omitted for clarity. In the space-filling model of **6**: gold, Au; green, the specified carbonyl carbon; pink, the specified carbons; blue, the specified hydrogen; grey, all other carbons; white, all other hydrogens. 'equiv.', equivalents.  $^t\text{Bu}$ , normal butyl;  $^i\text{Pr}$ , isopropyl.

redox-active metals (such as Rh, Ir, Ni, Ru and Fe) and have been carried out under harsh reaction conditions<sup>24,25</sup>, suggesting that the sterically unencumbered cationic IPrAu(I) complex might display a comparatively fast rate of coordination and subsequent oxidative addition.

We also examined the oxidative addition of IPrAu(I)SbF<sub>6</sub> with the more electron-rich 2,3,6,7-tetramethylbiphenylene (Me<sub>4</sub>-biphenylene). Full conversion to [IPrAu(III)(Me<sub>4</sub>-biphenyl)(H<sub>2</sub>O)][SbF<sub>6</sub>] (compound 7) was observed in 5 min at room temperature. When replacing AgSbF<sub>6</sub> with AgOTf, the oxidative addition was complete after 6 h, giving [IPrAu(III)(Me<sub>4</sub>-biphenyl)OTf] (compound 12). The longer reaction time required in the presence of the more coordinating anion is consistent with the hypothesis that coordination of biphenylene to a cationic coordinatively unsaturated Au(I) complex is the first step in the formation of 5. The thermodynamic driving force for this reaction comes from the cleavage of the strained biphenylene C–C bond (bond dissociation energy, BDE = 65.4 kcal mol<sup>−1</sup>) via oxidative addition to Au(I), yielding two Au–C(Ar) bonds<sup>24,25</sup>. This unprecedented C–C bond cleavage by the linear IPrAu(I) cationic complex represents a facile approach for generating strong Au–C(Ar) bonds en route to stable Au(III) complexes. Treatment of 5 with <sup>n</sup>Bu<sub>4</sub>NCl in dichloromethane resulted in the coordination of chloride and yielded IPrAu(III)(biphenyl)Cl (compound 3) as a pale yellow powder in 80% isolated yield after column chromatography (Fig. 2).

An X-ray crystal structure of 3 reveals a C<sub>s</sub>-symmetric distorted square-planar complex with one IPr carbene ligand, one chloride and two cyclometallation bonds from the biphenyl ligand defining the d<sup>8</sup> Au(III) geometry shown in Fig. 2. The Au–C bond distances in the gold–biphenyl of 2.028(12) Å and 2.046(11) Å, respectively, are shorter than the gold–carbene bond length of 2.117(11) Å (see Supplementary Information). These relatively short bond distances exhibit the strong bonding energy between high-oxidation-state Au(III) and its ligands. Consistent with our hypothesis, introduction of the cyclometallated biaryl ligand and one *N*-heterocyclic carbene ligand stabilized this high-oxidation-state Au(III) complex. The air- and moisture-stable complex 3 could be isolated and stored on the benchtop without any decomposition. Reaction of 3 with one equivalent of AgSbF<sub>6</sub> caused immediate precipitation of AgCl in CD<sub>2</sub>Cl<sub>2</sub> and the formation of 5 as observed by <sup>1</sup>H NMR. This feature allows 3 to be used as a stable precatalyst of cationic Au(III).

To gain more insight into the coordination chemistry of the cationic Au(III) complex with oxygen-based ligands, treatment of complex 5 with 1.1 equivalents of *N,N*-dimethylformamide (DMF) led to a substantial boost in stability, and allowed the isolation of [IPrAu(III)(biphenyl)(Me<sub>2</sub>NC(O)H)][SbF<sub>6</sub>] (compound 6) via coordination of a lone pair of electrons on the carbonyl oxygen. As shown by the crystal structure of complex 6 (Fig. 2), the IPr, biphenyl and DMF ligands enforce a distorted square-planar geometry around the Au(III) centre, with the oxygen atom of DMF at a distance of 2.140(3) Å from the centre. This finding suggests that the IPrAu(III)(biphenyl) cation exhibits a relatively hard, oxophilic Lewis-acidity, which is complementary to the soft Lewis acidity of cationic L–Au(I) complexes. This was further exemplified by measuring the Lewis acidities of 1 and 3 using the Gutmann–Beckett method (see Supplementary Information and Supplementary Fig. 3). The DMF in 6 is located in the pocket created by the IPr and biphenyl ligands. This environment is substantially more crowded than that observed in the linear IPrAuCl complex and effectively shields the DMF carbonyl group. The structural and electronic observations gained from this structure suggested a possible catalytic application of the IPrAu(III)(biphenyl) cation in selective 1,4-additions to  $\alpha,\beta$ -unsaturated aldehydes.

### Gold(III) complexes as selective Lewis-acid catalysts

Lewis-acid-catalysed 1,4-additions to  $\alpha,\beta$ -unsaturated aldehydes remain challenging because 1,2-additions generally predominate<sup>32–34</sup>. Yamamoto and colleagues have reported a unique strategy for Lewis-acid-promoted 1,4-addition to a  $\alpha,\beta$ -unsaturated aldehyde using aluminium reagents with sterically demanding ligands<sup>35</sup>. Although excellent 1,4-selectivity is obtained, stoichiometric amounts of Lewis acid appear to be necessary.

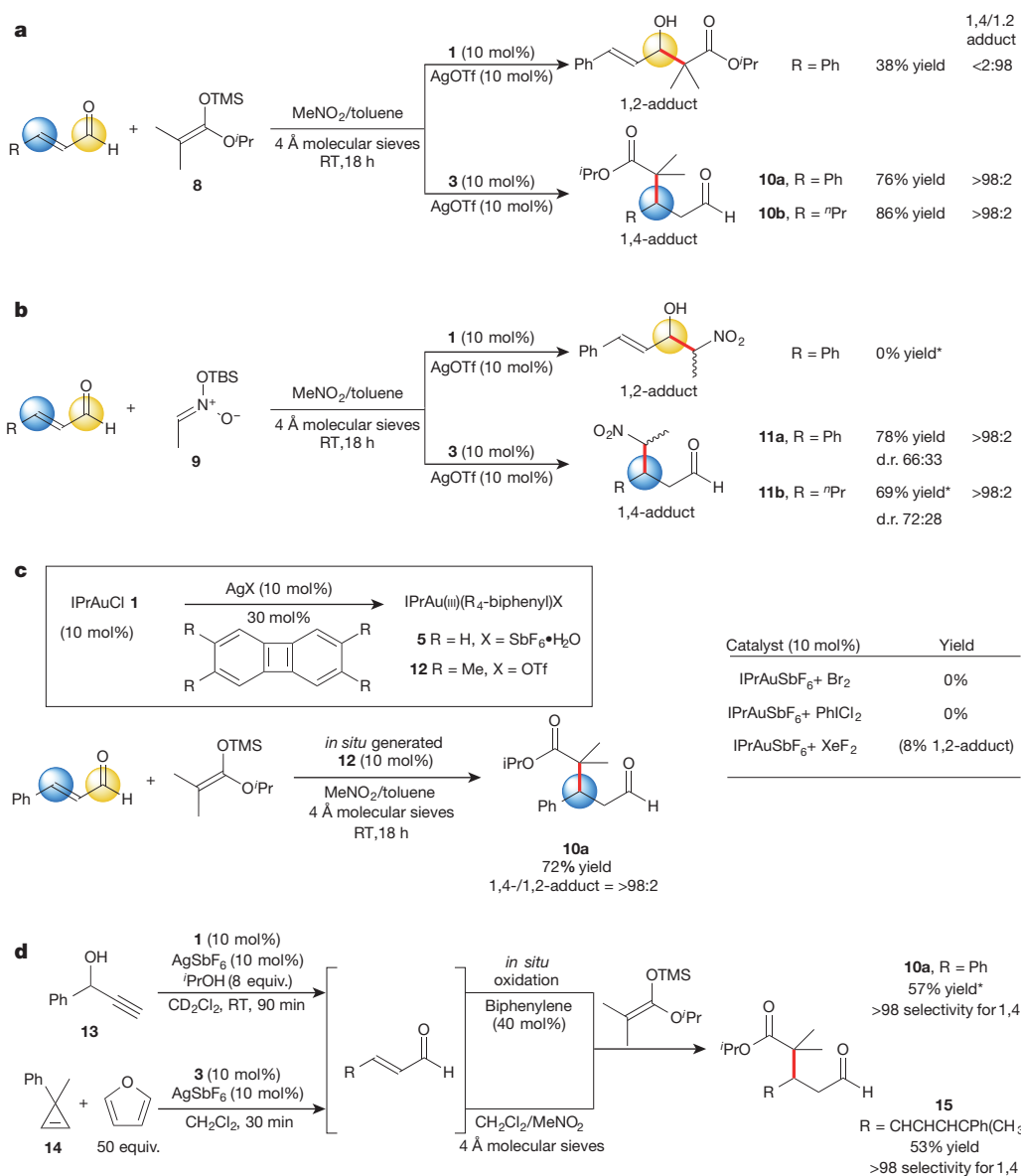
To obtain preliminary information regarding the reactivity of complex 5 prepared from 3 in Lewis-acid catalysis, we initiated our experimental efforts by studying the Mukaiyama–Michael reaction with cinnamaldehyde and ketene silyl acetal (compound 8) under ambient reaction conditions. As an initial result, the 1,4-adduct was obtained as the major product (1,4-/1,2-adduct = 75/25) in 98% NMR yield when catalysed by IPrAu(III)(biphenyl)(SbF<sub>6</sub>) (see Supplementary Information and Supplementary Table 2). Replacing AgSbF<sub>6</sub> with AgOTf gave 1,4-selectivity exclusively and the product was isolated in 76% yield (1,4-/1,2-adduct  $\geq$  98/2; Fig. 3a). With the optimized conditions in hand, silyl nitronate (compound 9) was examined as nucleophile instead of 8. The corresponding  $\gamma$ -nitro aldehyde was isolated in 78% yield with excellent selectivity (Fig. 3b). Employing *trans*-2-hexenal (compound 18a) as an aliphatic substrate also yielded the corresponding 1,4-adducts with good yield and excellent regioselectivity (Fig. 3a,b). Additionally, the catalyst loading could be decreased to 5 mol% without any loss of yield or selectivity: 10a was isolated in 77% yield with 1,4-/1,2-adduct  $>$  98/2 (see Supplementary Information, Supplementary Table 2). The cationic IPrAu(III)(biphenyl) complex was essential for 1,4-selectivity in these reactions; in control experiments, the corresponding 1,2-adduct was mainly obtained using IPrAu(I)(OTf) and no product was obtained with Ag(OTf) alone. To maximize the convenience and accessibility of our new catalyst, we examined the possibility of *in situ* generation of the IPrAu(III) cation via oxidative addition with biphenylene, followed by its utilization as a catalyst in one-pot reactions (Fig. 3c). Indeed, combining IPrAu(I)Cl, AgOTf, and Me<sub>4</sub>-biphenylene (30 mol%) led to the formation of the active catalyst IPrAu(III)(Me<sub>4</sub>-biphenyl)(OTf) (compound 12) within 1 h. Subsequent addition of cinnamaldehyde and ketene silyl acetal (8) led to the desired Mukaiyama–Michael adduct in 72% yield (1,4-/1,2-adduct  $\geq$  98/2). Notably, the reactions do not suffer from using an *in situ* generated catalyst as opposed to a preformed Au(III) complex. This flexibility illustrates the power of using biphenylene to generate the Au(III) catalyst: other oxidizing agents are completely incompatible with the sensitive ketene silyl acetal and aldehyde functionalities. As an illustration, replacement of biphenylene with Br<sub>2</sub> or PhICl<sub>2</sub> for the oxidation of IPrAu(I)SbF<sub>6</sub> yielded no 1,2-/1,4-adducts in the Mukaiyama–Michael reaction, and only 8% conversion to the 1,2-adduct was obtained with XeF<sub>2</sub> (Fig. 3c).

The ability to generate the Au(III) catalyst *in situ* under mild conditions suggested the possibility of performing successive Au(I)- and Au(III)-catalysed reactions in a single pot. Thus, we explored a one-pot tandem reaction starting from propargyl alcohol (compound 13 in Fig. 3d). First, the IPrAu(I)(SbF<sub>6</sub>) catalysed Meyer–Schuster rearrangement of propargyl alcohol (13) gave an unsaturated aldehyde. Subsequent oxidative addition of biphenylene afforded the IPrAu(III) cation 5, which catalysed the Mukaiyama–Michael addition of ketene silyl acetal (8) to provide the final product in 57% overall yield (1,4-/1,2-adduct  $\geq$  98/2). The ability to carry out two distinct reactions with catalysts of different oxidation state originating from a single precursor is a remarkable feature of the newly developed methods.

Furthermore, a successive Au(III)-catalysed ring-opening and Mukaiyama–Michael reaction was conducted. In the first step, 5 activated the ring-opening of the cyclopropene 14, which reacted with furan, to form a functionalized conjugated trienal (Fig. 3d)<sup>36</sup>. After removal of the excess furan, 5 catalysed the Mukaiyama–Michael addition of 8 to obtain the final product in 53% isolated yield ( $>$ 98% 1,4-selectivity).

To demonstrate the generality of this Au(III) catalyst for obtaining remote selectivity, we next performed 1,6-selective thiol addition and reduction reactions of  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes. We hypothesized that the bulky catalyst ‘aldehyde pocket’ should partially shield the proximal double bond, thus promoting nucleophilic addition at the  $\gamma,\delta$ -double bond. 1,6-additions to  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes are challenging and especially thiol additions and Hantzsch-ester-mediated reductions are known to proceed with low selectivity to provide a mixture of products<sup>37,38</sup>.





**Figure 3 | Examples of selective Au(III)-catalysed 1,4-additions.**

**a**, Mukaiyama–Michael addition. Tf, trifluoromethanesulfonate. **b**, Nitronate Michael-addition. d.r., diastereomeric ratio; TBS, *tert*-butyl dimethyl silyl.

For the thiol addition (Fig. 4a), 1,6-addition of naphthalene thiol to **16b** and subsequent oxidation was carried out. With the use of 10 mol% **3** and AgSbF<sub>6</sub>, the oxidized 1,6-addition product **17** was obtained with 64% yield and exclusive 1,6-selectivity in the presence of TBHP (*tert*-butyl hydroperoxide) as an oxidant. For the reduction reaction with Hantzsch ester (Fig. 4a), catalyst **5** also demonstrated excellent selectivity for the remote reduction of  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes **16** to give the  $\alpha,\beta$ -unsaturated aldehydes **18**, further showcasing the unique regioselectivity obtained with the Au(III) catalyst compared to traditional bulky Lewis- and Brønsted-acid catalysts (see Supplementary Information, Supplementary Table 5 and refs 32–34, for example).

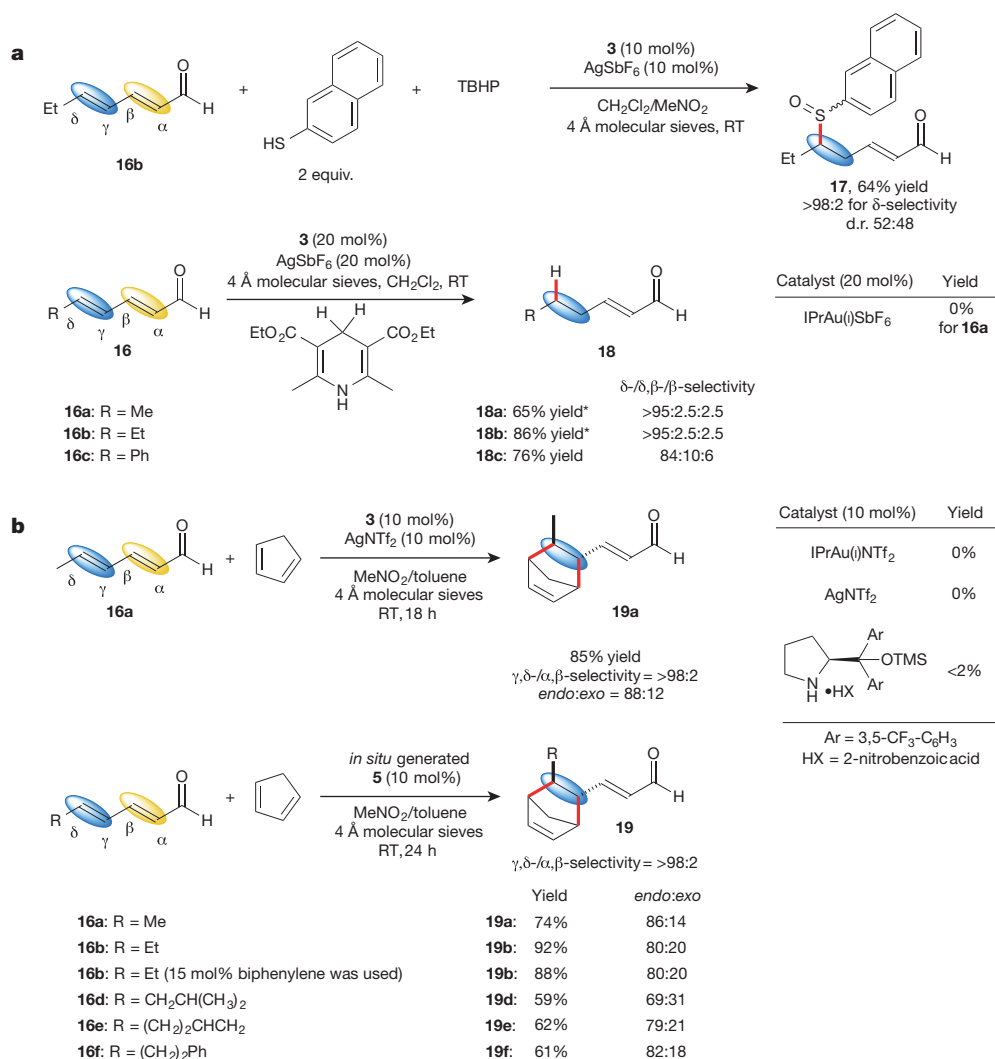
Moreover, we also performed the Diels–Alder reaction of 2,4-hexadienal (**16**) with cyclopentadiene<sup>39</sup>. Only the  $\gamma,\delta$ -functionalized product **19a** was obtained in 85% yield (diastereomeric ratio *endo:exo* = 88:12), employing IPrAu(III)(biphenyl)(NTf<sub>2</sub>) as catalyst (Fig. 4b).  $\gamma,\delta$ -selective Diels–Alder reactions are unprecedented because  $\alpha,\beta$ -selectivity normally predominates in such Lewis-acid-catalysed reactions<sup>40</sup>. Notably, neither IPrAu(I)(NTf<sub>2</sub>), Ag(NTf<sub>2</sub>) nor organocatalysts provided any product with **16a** and cyclopentadiene (see Supplementary Information and

**c**, *In situ* generation of IPrAu(III)(Me<sub>4</sub>-biphenyl) catalyst for

Mukaiyama–Michael addition. **d**, One-pot tandem Au(I)/Au(III)- and Au(III)/Au(III)-catalysed reactions. \*The yield is determined by NMR.

Supplementary Table 3), further illustrating the novel reactivity of this Au(III)-based catalytic system. Furthermore, *in situ* generation of Au(III) catalyst from Au(I) (Figs 3c and 4b) was also used for the Diels–Alder reaction. Several substrates were shown to provide products in good yield and excellent regioselectivity. The products were isolated in up to 92% yield exclusively as the  $\gamma,\delta$ -adducts, with an *endo:exo* ratio of up to 86:14, employing [IPrAu(III)(biphenyl)(H<sub>2</sub>O)]SbF<sub>6</sub> generated *in situ* from IPrAu(I)(SbF<sub>6</sub>) and biphenylene. To increase the practicality of the developed procedure we also showed that the amount of biphenylene oxidant can be lowered to 15 mol% (1.5 equivalents, as compared to Au(I)). Using these conditions, the Au(III) catalyst formed *in situ* produced product **19b** without a notable decline in yield and selectivity.

Whereas [2 + 2] cycloadditions of  $\alpha,\beta$ -unsaturated carbonyl compounds and allenes have been carried out with photochemical conditions<sup>41</sup>, there are no examples of Lewis-acid catalysis performing such reactions. Therefore, we next attempted to carry out the intramolecular [2 + 2] cycloaddition of the *cis*-unsaturated aldehyde-allene **20** (Fig. 5). When **3** with AgSbF<sub>6</sub> was employed, only the [2 + 2] cycloadduct was obtained in 70% yield (*cis:trans* = 89:11). Investigations showed that



**Figure 4** | Remote selectivity in Au(III)-catalysed additions to dienals. **a**,  $\delta$ -selective thiol addition and reduction reactions. Et, ethyl; cat, catalyst. **b**,  $\gamma,\delta$ -selective Diels–Alder reaction and *in situ* generation of the IPrAu(III)(biphenyl) catalyst for Diels–Alder reactions.

neither IPrAu(I)(SbF<sub>6</sub>), Ag(SbF<sub>6</sub>), MeAlCl<sub>2</sub>, *h* $\nu$  (irradiation by Hg vapour lamp) nor organocatalysts catalysed the [2 + 2] cycloaddition (see Supplementary Information and Supplementary Table 4); as such, this [2 + 2] cycloaddition is an unprecedented reaction unique to the developed Au(III) catalytic system.

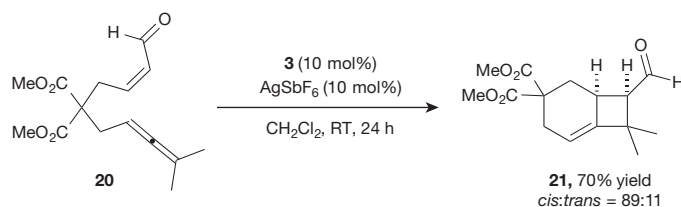
The solid-state structure of the key intermediate [IPrAu(III)(biphenyl)( $\eta^1$ -cinnamaldehyde)][SbF<sub>6</sub>] (compound **22**) was determined by X-ray crystallography, as illustrated in Fig. 6. This Au(III)–cinnamaldehyde complex displays a distorted square-planar geometry with a  $\eta^1$ -coordination cinnamaldehyde. The *s-trans*-cinnamaldehyde is buried in the pocket created by the IPr and biphenyl ligands. One of the 2,6-diisopropyl phenyl units is tilted away from the cinnamaldehyde plane owing to steric hindrance. As a result, the carbonyl moiety is effectively shielded from nucleophilic attack and thus the IPrAu(III)(biphenyl) cation exhibits

excellent remote selectivity as a catalyst for conjugate addition reactions of unsaturated aldehydes.

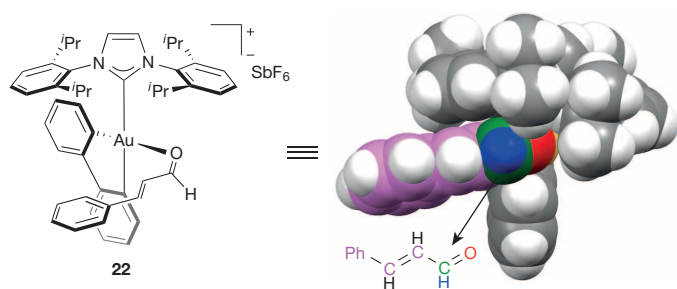
## Conclusion

Exploration of the unique catalytic abilities of high-valent late transition metals has been hampered by difficulties in accessing stable complexes with well controlled reactivity. Nowhere is this more evident than in the case of Au(III), which lags in its infancy compared to the ever-expanding field of Au(I) catalysis. In sharp contrast to previous methods for accessing Au(III), which rely on strong halogen-based oxidants, we have discovered that stable and catalytically active Au(III) complexes can be obtained by the mild oxidative addition of biphenylene to cationic IPrAu(I). Although oxidative additions with Au(I) have been previously viewed as kinetically challenging<sup>42,43</sup>, this carbon–carbon bond cleavage proceeds under surprisingly mild reaction conditions, even when compared to previously reported oxidative additions of biphenylene. The resulting IPrAu(III)(biphenyl) catalyst shows good reactivity as a hard Lewis-acid catalyst, which is complementary to the soft Lewis acidity exhibited by Au(I)-catalysts.

The catalytic reactivity is exemplified by six reactions of  $\alpha,\beta$ -unsaturated or  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes: Mukaiyama–Michael additions, nitronate Michael additions, a thiol addition, Hantzsch ester reductions, Diels–Alder reactions and a [2 + 2] cycloaddition all have good yield and excellent selectivity. Whereas often-used Au(III) catalysts like AuCl<sub>3</sub>



**Figure 5** | Au(III)-catalysed [2 + 2] cycloaddition of an allene-aldehyde.



**Figure 6 | A model for the obtained selectivity.** X-ray structure of the [IPrAu(III)(biphenyl)( $\eta^1$ -cinnamaldehyde)][SbF<sub>6</sub>] complex (**22**). The SbF<sub>6</sub><sup>−</sup> anion is omitted for clarity.

exhibit harsh, nonselective acidity<sup>44</sup>, these new catalysts possess intermediate reactivity and a sterically defined binding pocket. The ligand environment of the new catalysts, characterized by X-ray crystallography, imparts precise 1,4- over 1,2-selectivity in the case of the Michael additions,  $\delta$ -selectivity in the thiol addition and reduction reactions as well as  $\gamma,\delta$ -selectivity in the Diels–Alder reactions of dienals.

None of these reactions are feasible with Au(I)-based catalysts or traditional bulky Lewis- and Brønsted-acid catalysts (see Supplementary Information and Supplementary Table 5 and refs 32–34, for example). Furthermore, the catalysts can be generated *in situ* from commercially available reagents, thereby eliminating the need for specialized organometallic techniques. The mild nature of the oxidation from Au(I) to Au(III) even permits successive Au(I)- and Au(III)-catalysed reactions in a single reaction vessel. In other words, one can take a single precursor and switch between two different reaction manifolds by simple *in situ* catalyst modification. In view of all these features, the methods presented here should aid in unlocking the potential of high-oxidation-state gold catalysis.

Received 24 June; accepted 19 November 2014.

- Hickman, A. J. & Sanford, M. S. High-valent organometallic copper and palladium in catalysis. *Nature* **484**, 177–185 (2012).
- Furuya, T., Kamlet, A. S. & Ritter, T. Catalysis for fluorination and trifluoromethylation. *Nature* **473**, 470–477 (2011).
- Lee, E. *et al.* A fluoride-derived electrophilic late-stage fluorination reagent for PET imaging. *Science* **334**, 639–642 (2011).
- Furuya, T. *et al.* Mechanism of C–F reductive elimination from palladium(IV) fluorides. *J. Am. Chem. Soc.* **132**, 3793–3807 (2010).
- Whitfield, S. R. & Sanford, M. S. Reactivity of Pd(II) complexes with electrophilic chlorinating reagents: Isolation of Pd(IV) products and observation of C–Cl bond-forming reductive elimination. *J. Am. Chem. Soc.* **129**, 15142–15143 (2007).
- Alsters, P. L. *et al.* Rigid five- and six-membered C,N,N'-bound aryl-, benzyl-, and alkylorganopalladium complexes:  $sp^2$  vs.  $sp^3$  carbon-hydrogen activation during cyclopalladation and palladium(IV) intermediates in oxidative addition reactions with dihalogens and alkyl halides. *Organometallics* **12**, 1831–1844 (1993).
- Wang, Y.-M., Lackner, A. D. & Toste, F. D. Development of catalysts and ligands for enantioselective gold catalysis. *Acc. Chem. Res.* **47**, 889–901 (2014).
- Krause, N. & Winter, C. Gold-catalyzed nucleophilic cyclization of functionalized allenes: a powerful access to carbo- and heterocycles. *Chem. Rev.* **111**, 1994–2009 (2011).
- Corma, A., Leyva-Pérez, A. & Sabater, M. J. Gold-catalyzed carbon-heteroatom bond-forming reactions. *Chem. Rev.* **111**, 1657–1712 (2011).
- Hamilton, G. L., Kang, E. J., Mba, M. & Toste, F. D. A powerful chiral counterion strategy for asymmetric transition metal catalysis. *Science* **317**, 496–499 (2007).
- Gorin, D. J. & Toste, F. D. Relativistic effects in homogeneous gold catalysis. *Nature* **446**, 395–403 (2007).
- Schmidbaur, H. & Schier, A. Gold(III) compounds for homogeneous catalysis: preparation, reaction conditions, and scope of application. *Arabian J. Sci. Eng.* **37**, 1187–1225 (2012).
- Oliver-Meseguer, J. *et al.* Small gold clusters formed in solution give reaction turnover numbers of  $10^7$  at room temperature. *Science* **338**, 1452–1455 (2012).
- Leyva-Pérez, A. & Corma, A. Similarities and differences between the “relativistic” triad gold, platinum, and mercury in catalysis. *Angew. Chem. Int. Ed.* **51**, 614–635 (2012).
- Gaillard, S. *et al.* Synthetic and structural studies of [AuCl<sub>3</sub>(NHC)] complexes. *Organometallics* **29**, 394–402 (2010).
- de Frémont, P., Singh, R., Stevens, E. D., Petersen, J. L. & Nolan, S. P. Synthesis, characterization and reactivity of *N*-heterocyclic carbene gold(III) complexes. *Organometallics* **26**, 1376–1385 (2007).
- Hashmi, A. S. K., Blanco, M. C., Fischer, D. & Bats, J. W. Gold catalysis: evidence for the *in situ* reduction of gold(III) during the cyclization of allenyl carbinols. *Eur. J. Org. Chem.* 1387–1389 (2006).

- Wolf, W. J., Winston, M. S. & Toste, F. D. Exceptionally fast carbon–carbon bond reductive elimination from gold(III). *Nature Chem.* **6**, 159–164 (2013).
- Vicente, J., Bermúdez, M.-D., Carrión, F.-J. & Jones, P. G. Synthesis and reactivity of some nitroaryl complexes of Hg<sup>II</sup> and Au<sup>III</sup> — synthesis of a substituted biphenyl by C–C coupling-crystal structure of [Hg(C<sub>6</sub>H<sub>4</sub>N<sub>2</sub>O<sub>2</sub>-3,OnBu-6)<sub>2</sub>]. *Chem. Ber.* **129**, 1395–1399 (1996).
- Roşca, D.-A., Smith, D. A., Hughes, D. L. & Bochmann, M. A thermally stable gold(III) hydride: synthesis, reactivity, and reductive condensation as a route to gold(II) complexes. *Angew. Chem. Int. Ed.* **51**, 10643–10646 (2012).
- Hashmi, A. S. K. Fire and ice: a gold(III) monohydride. *Angew. Chem. Int. Ed.* **51**, 12935–12936 (2012).
- Usón, R., Vicente, J., Cirac, J. A. & Chicote, M. T. Synthesis and reactivity of dibenzometalole complexes of gold(III) and platinum(II). *J. Organomet. Chem.* **198**, 105–112 (1980).
- Pyykkö, P. & Runeberg, N. Comparative theoretical study of *N*-heterocyclic carbenes and other ligands bound to Au<sup>I</sup>. *Chem. Asian J.* **1**, 623–628 (2006).
- Jones, W. D. Mechanistic studies of transition metal-mediated C–C bond activation. *Top. Curr. Chem.* **346**, 1–31 (2013).
- Perthuisot, C. *et al.* Cleavage of the carbon–carbon bond in biphenylene using transition metals. *J. Mol. Catal. Chem.* **189**, 157–168 (2002).
- Darmon, J. M. *et al.* Oxidative addition of carbon–carbon bonds with a redox-active bis(imino)pyridine iron complex. *J. Am. Chem. Soc.* **134**, 17125–17137 (2012).
- Levin, L. D. & Toste, F. D. Gold-catalyzed allylation of aryl boronic acids: accessing cross-coupling reactivity with gold. *Angew. Chem. Int. Ed.* **53**, 6211–6215 (2014).
- Fackler, J. P. Jr. Metal-metal bond formation in the oxidative addition to dinuclear gold(I) species. Implications from dinuclear and trinuclear gold chemistry for the oxidative addition process generally. *Polyhedron* **16**, 1–17 (1997).
- Guenther, J. *et al.* Activation of aryl halides at gold(I): practical synthesis of (P,C) cyclometalated gold(III) complexes. *J. Am. Chem. Soc.* **136**, 1778–1781 (2014).
- Jost, M. *et al.* Facile oxidative addition of aryl iodides to gold(I) by ligand design: bending turns on reactivity. *J. Am. Chem. Soc.* **136**, 14654–14657 (2014).
- Weber, S. G., Rominger, F. & Straub, B. F. Isolated silver intermediate of gold precatalyst activation. *Eur. J. Inorg. Chem.* 2863–2867 (2012).
- North, M., Usanov, D. L. & Young, C. Lewis acid catalyzed asymmetric cyanohydrin synthesis. *Chem. Rev.* **108**, 5146–5226 (2008).
- Yamamoto, H. (ed.) *Lewis Acids in Organic Synthesis* 1–995 (Wiley, 2000).
- Mahrwald, R. Diastereoselection in Lewis-acid-mediated aldol reactions. *Chem. Rev.* **99**, 1095–1120 (1999).
- Maruoka, K., Imoto, H., Saito, S. & Yamamoto, H. Virtually complete blocking of  $\alpha,\beta$ -unsaturated aldehyde carbonyls by complexation with aluminum tris(2,6-diphenylphenoxide). *J. Am. Chem. Soc.* **116**, 4131–4132 (1994).
- Hadfield, M. S. & Lee, A.-L. Gold(I)-catalyzed synthesis of conjugated trienes. *Chem. Commun.* **47**, 1333–1335 (2011).
- Akagawa, K., Nishi, N., Sen, J. & Kudo, K. Peptide-catalyzed consecutive 1,6- and 1,4-additions of thiols to  $\alpha,\beta,\gamma,\delta$ -unsaturated aldehydes. *Org. Biomol. Chem.* **12**, 3581–3585 (2014).
- Akagawa, K., Sen, J. & Kudo, K. Peptide-catalyzed regio- and enantioselective reduction of  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes. *Angew. Chem. Int. Ed.* **52**, 11585–11588 (2013).
- Ahrendt, K. A., Borths, C. J. & MacMillan, D. W. C. New strategies for organic synthesis: the first highly enantioselective organocatalytic Diels–Alder reaction. *J. Am. Chem. Soc.* **122**, 4243–4244 (2000).
- Hayashi, Y., Okamura, D., Umehira, S. & Uchimaru, T. Organocatalytic 1,4-addition reaction of  $\alpha,\beta,\gamma,\delta$ -diunsaturated aldehydes versus 1,6-addition reaction. *ChemCatChem* **4**, 959–962 (2012).
- Alcaide, B., Almendros, P. & Arancillo, C. Exploiting [2 + 2] cycloaddition chemistry: achievements with allenes. *Chem. Soc. Rev.* **39**, 783–816 (2010).
- Hashmi, A. S. K. *et al.* On homogeneous gold/palladium catalytic systems. *Adv. Synth. Catal.* **354**, 133–147 (2012).
- Livendahl, M., Goehry, C., Maseras, F. & Echavarren, A. M. Rationale for the sluggish oxidative addition of aryl halides to Au(I). *J. Chem. Commun.* **50**, 1533–1536 (2014).
- Hashmi, A. S. K., Frost, T. M. & Bats, J. W. Highly selective gold-catalyzed arene synthesis. *J. Am. Chem. Soc.* **122**, 11553–11554 (2000).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We gratefully acknowledge the NIHGMs (R01 GM073932) for financial support. C.-Y.W. thanks the Taiwan National Science Council for a postdoctoral fellowship (2011–2012). T.H. thanks the Uehara Memorial Foundation for a postdoctoral fellowship. C.B.J. is grateful to the Lundbeck Foundation for a postdoctoral fellowship. We thank A. DiPasquale (at the College of Chemistry X-ray Crystallography Facility of the University of California, Berkeley) for X-ray crystallographic data collection and we acknowledge support from the NIH Shared Instrumentation Grant S10-RR027172. We thank H.-J. Liu for his generous donation of the biphenylene.

**Author Contributions** C.-Y.W. initiated and developed the organometallic study. C.-Y.W. and T.H. developed the Lewis-acid catalysis. C.-Y.W., T.H. and C.B.J. optimized the Lewis-acid catalysis study. C.-Y.W., T.H. and C.B.J. performed the experiments. C.-Y.W., T.H. and C.B.J. and F.D.T. wrote the manuscript.

**Author Information** X-ray crystallographic data have been deposited in the Cambridge Crystallographic Data Centre database (<http://www.ccdc.cam.ac.uk/>) under code CCDC 1002525–1002527. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.D.T. (fdtoste@berkeley.edu).



# A new antibiotic kills pathogens without detectable resistance

Losee L. Ling<sup>1\*</sup>, Tanja Schneider<sup>2,3\*</sup>, Aaron J. Peoples<sup>1</sup>, Amy L. Spoering<sup>1</sup>, Ina Engels<sup>2,3</sup>, Brian P. Conlon<sup>4</sup>, Anna Mueller<sup>2,3</sup>, Till F. Schäberle<sup>3,5</sup>, Dallas E. Hughes<sup>1</sup>, Slava Epstein<sup>6</sup>, Michael Jones<sup>7</sup>, Linos Lazarides<sup>7</sup>, Victoria A. Steadman<sup>7</sup>, Douglas R. Cohen<sup>1</sup>, Cintia R. Felix<sup>1</sup>, K. Ashley Fetterman<sup>1</sup>, William P. Millett<sup>1</sup>, Anthony G. Nitti<sup>1</sup>, Ashley M. Zullo<sup>1</sup>, Chao Chen<sup>4</sup> & Kim Lewis<sup>4</sup>

**Antibiotic resistance is spreading faster than the introduction of new compounds into clinical practice, causing a public health crisis. Most antibiotics were produced by screening soil microorganisms, but this limited resource of cultivable bacteria was overmined by the 1960s. Synthetic approaches to produce antibiotics have been unable to replace this platform. Uncultured bacteria make up approximately 99% of all species in external environments, and are an untapped source of new antibiotics. We developed several methods to grow uncultured organisms by cultivation *in situ* or by using specific growth factors. Here we report a new antibiotic that we term teixobactin, discovered in a screen of uncultured bacteria. Teixobactin inhibits cell wall synthesis by binding to a highly conserved motif of lipid II (precursor of peptidoglycan) and lipid III (precursor of cell wall teichoic acid). We did not obtain any mutants of *Staphylococcus aureus* or *Mycobacterium tuberculosis* resistant to teixobactin. The properties of this compound suggest a path towards developing antibiotics that are likely to avoid development of resistance.**

Widespread introduction of antibiotics in the 1940s, beginning with penicillin<sup>1,2</sup> and streptomycin<sup>3</sup>, transformed medicine, providing effective cures for the most prevalent diseases of the time. Resistance development limits the useful lifespan of antibiotics and results in the requirement for a constant introduction of new compounds<sup>4,5</sup>. However, antimicrobial drug discovery is uniquely difficult<sup>6</sup>, primarily due to poor penetration of compounds into bacterial cells. Natural products evolved to breach the penetration barriers of target bacteria, and most antibiotics introduced into the clinic were discovered by screening cultivable soil microorganisms. Overmining of this limited resource by the 1960s brought an end to the initial era of antibiotic discovery<sup>7</sup>. Synthetic approaches were unable to replace natural products<sup>8</sup>.

Approximately 99% of all species in external environments are uncultured (do not grow under laboratory conditions), and are a promising source of new antibiotics<sup>8</sup>. We developed several methods to grow uncultured organisms by cultivation in their natural environment<sup>9,10</sup>, or by using specific growth factors such as iron-chelating siderophores<sup>11</sup>. Uncultured organisms have recently been reported to produce interesting compounds with new structures/modes of action—lassomycin, an inhibitor of the essential mycobacterial protease ClpP1P2C1 (ref. 12); and diverse secondary metabolites present in a marine sponge *Theonella swinhoei* which are actually made by an uncultured symbiotic *Entotheonella* sp.<sup>13</sup>.

Here we report the discovery of a new cell wall inhibitor, teixobactin, from a screen of uncultured bacteria grown in diffusion chambers *in situ*.

## Identification of teixobactin

A multichannel device, the iChip<sup>10</sup>, was used to simultaneously isolate and grow uncultured bacteria. A sample of soil is diluted so that approximately one bacterial cell is delivered to a given channel, after which the device is covered with two semi-permeable membranes and placed back in the soil (Extended Data Fig. 1). Diffusion of nutrients and growth

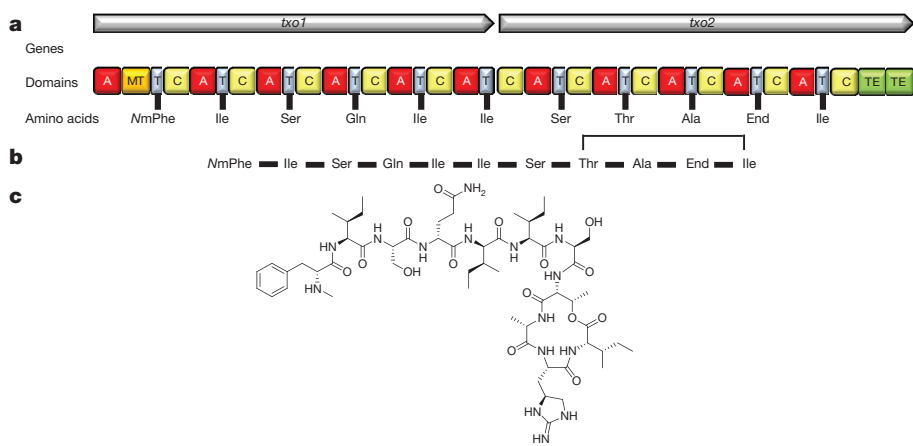
factors through the chambers enables growth of uncultured bacteria in their natural environment. The growth recovery by this method approaches 50%, as compared to 1% of cells from soil that will grow on a nutrient Petri dish<sup>10</sup>. Once a colony is produced, a substantial number of uncultured isolates are able to grow *in vitro*<sup>14</sup>. Extracts from 10,000 isolates obtained by growth in iChips were screened for antimicrobial activity on plates overlaid with *S. aureus*. An extract from a new species of  $\beta$ -proteobacteria provisionally named *Eleftheria terrae* showed good activity. The genome of *E. terrae* was sequenced (Supplementary Discussion). Based on 16S rDNA and *in silico* DNA/DNA hybridization, this organism belongs to a new genus related to *Aquabacteria* (Extended Data Fig. 2, Supplementary Discussion). This group of Gram-negative organisms is not known to produce antibiotics. A partially purified active fraction contained a compound with a molecular mass of 1,242 Da determined by mass spectrometry, which was not reported in available databases. The compound was isolated and a complete stereochemical assignment has been made based on NMR and advanced Marfey's analysis (Fig. 1, Extended Data Figs 3 and 4 and Supplementary Discussion). This molecule, which we named teixobactin, is an unusual depsipeptide which contains enduracididine, methylphenylalanine, and four D-amino acids. The biosynthetic gene cluster (GenBank accession number KP006601) was identified using a homology search (Supplementary Discussion). It consists of two large non-ribosomal peptide synthetase (NRPS)-coding genes, which we named *txo1* and *txo2*, respectively (Fig. 1). In accordance with the co-linearity rule, 11 modules are encoded. The *in silico* predicted adenylation domain specificity perfectly matches the amino acid order of teixobactin (Fig. 1), and allowed us to predict the biosynthetic pathway (Extended Data Fig. 5).

## Resistance and mechanism of action

Teixobactin had excellent activity against Gram-positive pathogens, including drug-resistant strains (Table 1 and Extended Data Table 1).

<sup>1</sup>NovoBiotic Pharmaceuticals, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Institute of Medical Microbiology, Immunology and Parasitology—Pharmaceutical Microbiology Section, University of Bonn, Bonn 53115, Germany. <sup>3</sup>German Centre for Infection Research (DZIF), Partner Site Bonn-Cologne, 53115 Bonn, Germany. <sup>4</sup>Antimicrobial Discovery Center, Northeastern University, Department of Biology, Boston, Massachusetts 02115, USA. <sup>5</sup>Institute for Pharmaceutical Biology, University of Bonn, Bonn 53115, Germany. <sup>6</sup>Department of Biology, Northeastern University, Boston, Massachusetts 02115, USA. <sup>7</sup>Selcia, Ongar, Essex CM5 0GS, UK.

\*These authors contributed equally to this work.



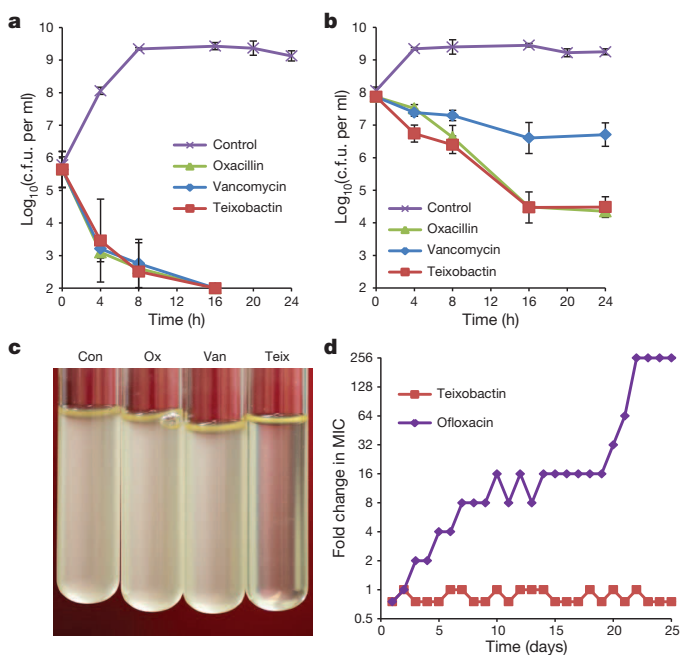
**Figure 1 | The structure of teixobactin and the predicted biosynthetic gene cluster.** **a**, The two NRPS genes, the catalytic domains they encode, and the amino acids incorporated by the respective modules. Domains: A, adenylation; C, condensation; MT, methylation (of phenylalanine); T, thiolation (carrier); and TE, thioesterase (Ile-Thr ring closure). NmPhe, *N*-methylated phenylalanine. **b**, Schematic structure of teixobactin. The *N*-methylation of the first phenylalanine is catalysed by the methyltransferase domain in module 1. The ring closure between the last isoleucine and threonine is catalysed by the thioesterase domains during molecule off-loading, resulting in teixobactin. **c**, Teixobactin structure.

Potency against most species, including difficult-to-treat enterococci and *M. tuberculosis* was below  $1 \mu\text{g ml}^{-1}$ . Teixobactin was exceptionally active against *Clostridium difficile* and *Bacillus anthracis* (minimal inhibitory concentration (MIC) of 5 and  $20 \text{ ng ml}^{-1}$ , respectively). Teixobactin had excellent bactericidal activity against *S. aureus* (Fig. 2a), was superior to vancomycin in killing late exponential phase populations (Fig. 2b), and retained bactericidal activity against intermediate resistance *S. aureus* (VISA) (Extended Data Fig. 6a). Note that frequent clinical failure in patients with *S. aureus* MRSA treated with vancomycin has been linked to the poor bactericidal activity of this compound<sup>15,16</sup>. Teixobactin was ineffective against most Gram-negative bacteria, but showed good activity against a strain of *E. coli* *asmB1* with a defective outer membrane permeability barrier (Table 1).

We were unable to obtain mutants of *S. aureus* or *M. tuberculosis* resistant to teixobactin even when plating on media with a low dose ( $4 \times \text{MIC}$ ) of the compound. Serial passage of *S. aureus* in the presence of sub-MIC levels of teixobactin over a period of 27 days failed to produce resistant mutants as well (Fig. 2d, Supplementary Discussion). This usually points to a non-specific mode of action, with accompanying toxicity. However, teixobactin had no toxicity against mammalian NIH/3T3 and HepG2 cells at  $100 \mu\text{g ml}^{-1}$  (the highest dose tested). The compound showed no haemolytic activity and did not bind DNA. In order to determine specificity of action of teixobactin, we examined its effect on the rate of label incorporation into the major biosynthetic pathways of *S. aureus*. Teixobactin strongly inhibited synthesis of peptidoglycan, but had virtually no effect on label incorporation into DNA,

RNA and protein (Fig. 3a). This suggested that teixobactin is a new peptidoglycan synthesis inhibitor.

Resistance has not developed to this compound, suggesting that the target is not a protein. The essential lack of resistance development through mutations has been described for vancomycin which binds lipid II, the precursor of peptidoglycan. We reasoned that teixobactin could be acting against the same target. Treatment of whole cells of *S. aureus* with teixobactin ( $1-5 \times \text{MIC}$ ) resulted in significant accumulation of the soluble cell wall precursor undecaprenyl-*N*-acetylmuramic acid-pentapeptide (UDP-MurNAc-pentapeptide), similar to the vancomycin-treated control cells (Fig. 3b), showing that one of the membrane-associated steps of peptidoglycan biosynthesis is blocked. Teixobactin inhibited peptidoglycan biosynthesis reactions *in vitro* in a dose-dependent manner with either lipid I, lipid II or undecaprenyl-pyrophosphate (Fig. 3c) as a substrate. Quantitative analysis of the

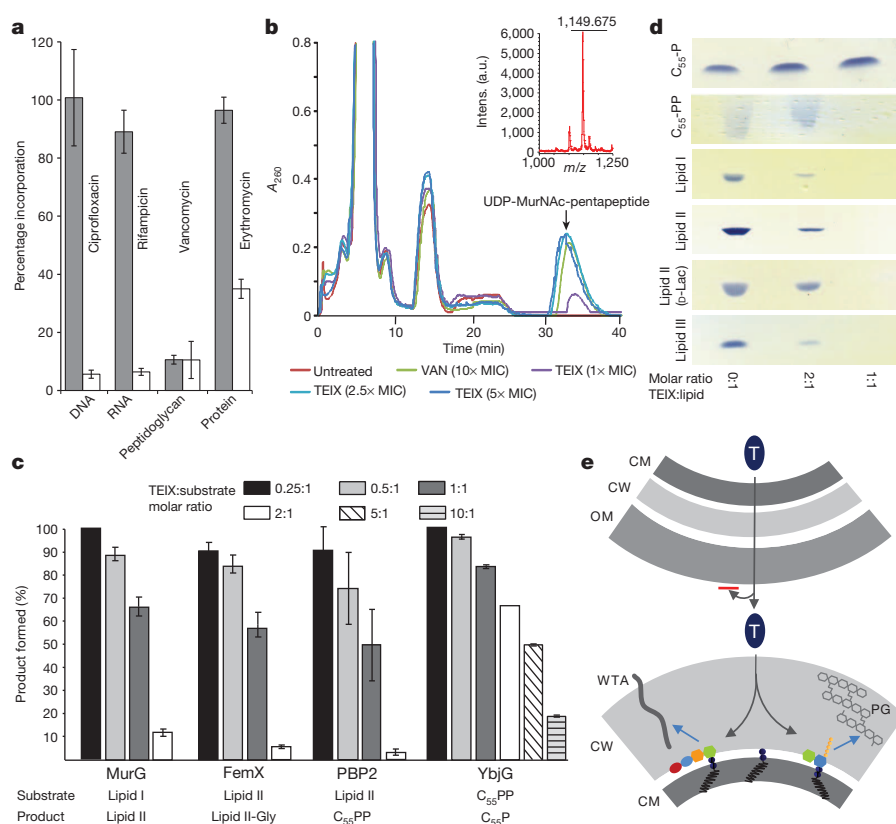


**Figure 2 | Time-dependent killing of pathogens by teixobactin.** **a, b,** *S. aureus* were grown to early (**a**), and late (**b**) exponential phase and challenged with antibiotics. Data are representative of 3 independent experiments  $\pm$  s.d. **c,** Teixobactin treatment resulted in lysis. The figure is representative of 3 independent experiments. **d,** Resistance acquisition during serial passaging in the presence of sub-MIC levels of antimicrobials. The y axis is the highest concentration the cells grew in during passaging. For ofloxacin,  $256 \times$  MIC was the highest concentration tested. The figure is representative of 3 independent experiments.

### Table 1 | Activity of teixobactin against pathogenic microorganisms

Organism and genotype	Teixobactin MIC ( $\mu\text{g ml}^{-1}$ )
<i>S. aureus</i> (MSSA)	0.25
<i>S. aureus</i> + 10% serum	0.25
<i>S. aureus</i> (MRSA)	0.25
<i>Enterococcus faecalis</i> (VRE)	0.5
<i>Enterococcus faecium</i> (VRE)	0.5
<i>Streptococcus pneumoniae</i> (penicillin <sup>R</sup> )	$\leq 0.03$
<i>Streptococcus pyogenes</i>	0.06
<i>Streptococcus agalactiae</i>	0.12
Viridans group streptococci	0.12
<i>B. anthracis</i>	$\leq 0.06$
<i>Clostridium difficile</i>	0.005
<i>Propionibacterium acnes</i>	0.08
<i>M. tuberculosis</i> H37Rv	0.125
<i>Haemophilus influenzae</i>	4
<i>Moraxella catarrhalis</i>	2
<i>Escherichia coli</i>	25
<i>Escherichia coli</i> (asmB1)	2.5
<i>Pseudomonas aeruginosa</i>	>32
<i>Klebsiella pneumoniae</i>	>32

The MIC was determined by broth microdilution. MSSA, methicillin-sensitive *S. aureus*; VRE, vancomycin-resistant enterococci.



**Figure 3 | Teixobactin binds to cell wall precursors.** **a**, Impact of teixobactin (TEIX) on macromolecular biosyntheses in *S. aureus*. Incorporation of  $^3\text{H}$ -thymidine (DNA),  $^3\text{H}$ -uridine (RNA),  $^3\text{H}$ -leucine (protein), and  $^3\text{H}$ -glucosamine (peptidoglycan) was determined in cells treated with teixobactin at  $1 \times \text{MIC}$  (grey bars). Ciprofloxacin ( $8 \times \text{MIC}$ ), rifampicin ( $4 \times \text{MIC}$ ), vancomycin ( $2 \times \text{MIC}$ ) and erythromycin ( $2 \times \text{MIC}$ ) were used as controls (white bars). Data are means of 4 independent experiments  $\pm$  s.d. **b**, Intracellular accumulation of the cell wall precursor UDP-MurNAc-pentapeptide after treatment of *S. aureus* with teixobactin. Untreated and vancomycin (VAN)-treated ( $10 \times \text{MIC}$ ) cells were used as controls. UDP-MurNAc-pentapeptide was identified by mass spectrometry as indicated by the peak at  $m/z$  1,149.675. The experiment is representative of 3 independent experiments. **c**, The effect of teixobactin on precursor consuming reactions. Experiments were performed in 3 biological replicates and data are presented as mean  $\pm$  s.d. **d**, Complex formation of teixobactin with purified cell wall precursors. Binding of teixobactin is indicated by a reduction of the amount of lipid intermediates (visible on the thin-layer chromatogram). The figure is representative of two independent experiments. **e**, A model of teixobactin targeting and resistance. The teixobactin producer is a Gram-negative bacterium protected from this compound by exporting it across the outer membrane permeability barrier (upper panel). In target Gram-positive organisms lacking an outer membrane, the targets are readily accessible on the outside where teixobactin binds precursors of peptidoglycan (PG) and WTA. CM, cytoplasmic membrane; CW, cell wall; OM, outer membrane; T, teixobactin.

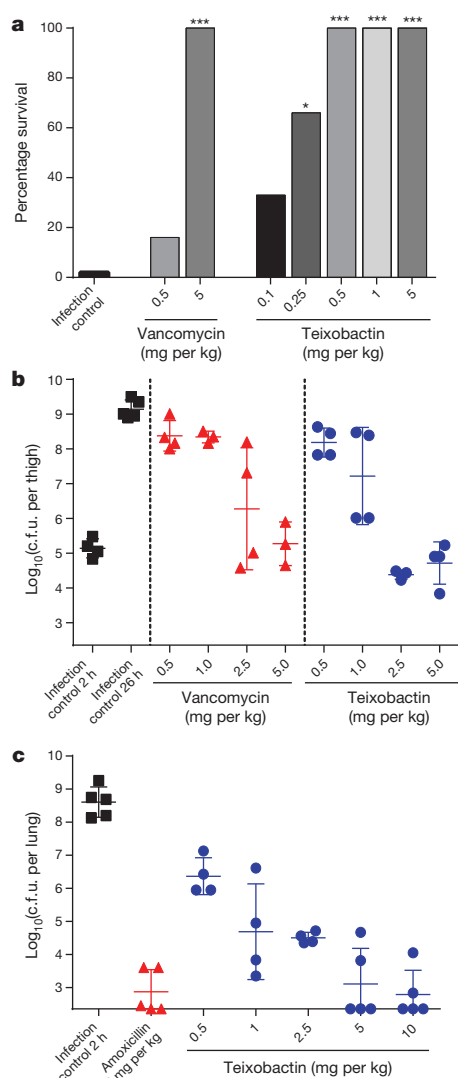
MurG-, FemX-, and PBP2-catalysed reactions using radiolabelled substrates, showed an almost complete inhibition at a twofold molar excess of teixobactin with respect to the lipid substrate (Fig. 3c). The addition of purified lipid II prevented teixobactin from inhibiting growth of *S. aureus* (Extended Data Table 2). These experiments showed that teixobactin specifically interacts with the peptidoglycan precursor, rather than interfering with the activity of one of the enzymes. In order to evaluate the minimal motif required for high affinity binding of teixobactin, the direct interaction with several undecaprenyl-coupled cell envelope precursors was investigated. Purified precursors were incubated with teixobactin at different molar ratios, followed by extraction and subsequent thin-layer chromatography analysis (Fig. 3d). In agreement with the results obtained from the *in vitro* experiments, lipid I and lipid II were fully trapped in a stable complex that prevented extraction of the lipid from the reaction mixture in the presence of a twofold molar excess of the antibiotic, leading to the formation of a 2:1 stoichiometric complex. Teixobactin was active against vancomycin-resistant enterococci that have modified lipid II (lipid II-D-Ala-D-Lac or lipid II-D-Ala-D-Ser instead of lipid II-D-Ala-D-Ala)<sup>17–19</sup>. This suggested that, unlike vancomycin, teixobactin is able to bind to these modified forms of lipid II. Indeed, teixobactin bound to lipid II-D-Ala-D-Lac and lipid II-D-Ala-D-Ser (Extended Data Fig. 6b). Moreover, teixobactin efficiently bound to the wall teichoic acid (WTA) precursor undecaprenyl-PP-GlcNAc (lipid III). Although WTA is not essential *per se*, inhibition of late membrane-bound WTA biosynthesis steps is lethal due to accumulation of toxic intermediates<sup>20</sup>. Furthermore, teichoic acids anchor autolysins, preventing uncontrolled hydrolysis of peptidoglycan<sup>21</sup>. Inhibition of teichoic acid synthesis by teixobactin would help liberate autolysins,

contributing to the excellent lytic and killing activity of this antibiotic. Teixobactin was also able to bind undecaprenyl-pyrophosphate, but not undecaprenyl-phosphate (Fig. 3d and Extended Data Table 2). Although teixobactin efficiently binds lipid I *in vitro*, this is probably less significant for antimicrobial activity, as this is the intracellular form of the precursor, unlike surface-exposed lipid II and the undecaprenyl-PP-GlcNAc WTA precursors (Fig. 3e and Extended Data Fig. 7). Binding to the target primarily relies on the interaction of the antibiotic with the pyrophosphate moiety, and the first sugar moiety attached to the lipid carrier, as higher concentrations of teixobactin were required to completely inhibit the YbjG-catalysed monophosphorylation of undecaprenyl-pyrophosphate, involved in the recycling process of the essential lipid carrier (Fig. 3c and Extended Data Fig. 7). Corroborating this result, a tenfold higher concentration of undecaprenyl-pyrophosphate was required to antagonize the antimicrobial activity of teixobactin compared to lipid II (Extended Data Table 2). The exact nature of this first sugar is therefore not important, explaining why teixobactin is active against *M. tuberculosis*, where it probably binds to decaprenyl-coupled lipid intermediates of peptidoglycan and arabinogalactan. Teixobactin is also likely to bind to prenyl-PP-sugar intermediates of capsular polysaccharide biosynthesis which is important for virulence in staphylococci<sup>22</sup> and whose inhibition of biosynthesis is lethal in streptococci<sup>23</sup>.

### In vivo efficacy

Given the attractive mode of action of this compound, we investigated its potential as a therapeutic. The compound retained its potency in the presence of serum, was stable, and had good microsomal stability and low toxicity (Supplementary Discussion). The pharmacokinetic





**Figure 4 | Teixobactin is efficacious in three mouse models of infection.** **a**, Single dose treatment (i.v., 1 h post-infection, 6 mice per group) with teixobactin and vancomycin in septicemia protection model using MRSA. Survival is depicted 48 h after infection. **b**, Single dose (i.v., 2 h post-infection, 4 mice per group) treatment with teixobactin and vancomycin in neutropenic mouse thigh infection model using MRSA. For drug-treated animals, thigh colony-forming units (c.f.u.) were determined at 26 h post-infection. For controls, c.f.u. in thighs were determined at 2 h and 26 h post-infection. **c**, Two dose treatment, 5 mice per group, with teixobactin (i.v., 24 h and 36 h post-infection) and single dose treatment with amoxicillin (subcutaneous, 24 h post-infection) in immunocompetent lung infection model using *S. pneumoniae*. Lung c.f.u. were determined at 48 h post-infection. The c.f.u. from each mouse are plotted as individual points and error bars represent the deviation within an experimental group. \* $P < 0.05$ , \*\*\* $P < 0.001$  (determined by non-parametric log-rank test).

parameters determined after i.v. injection of a single 20 mg per kg dose in mice were favourable, as the level of compound in serum was maintained above the MIC for 4 h (Extended Data Fig. 8). An animal efficacy study was then performed in a mouse septicemia model. Mice were infected intraperitoneally with methicillin-resistant *S. aureus* (MRSA) at a dose that leads to 90% of death. One hour post-infection, teixobactin was introduced i.v. at single doses ranging from 1 to 20 mg per kg. All treated animals survived (Fig. 4a), and in a subsequent experiment the PD<sub>50</sub> (protective dose at which half of the animals survive) was determined to be 0.2 mg per kg, which compares favourably to the 2.75 mg per kg PD<sub>50</sub> of vancomycin, the main antibiotic used to treat MRSA. Teixobactin was then tested in a thigh model of infection with *S. aureus*,

and showed good efficacy as well (Fig. 4b). Teixobactin was also highly efficacious in mice infected with *Streptococcus pneumoniae*, causing a 6 log<sub>10</sub> reduction of c.f.u. in lungs (Fig. 4c).

## Discussion

This study, as well as previous work<sup>12,13,24</sup> suggests that new organisms such as uncultured bacteria are likely to harbour new antimicrobials<sup>25</sup>. This is consistent with resistance mechanisms in soil bacteria being stratified by phylogeny, with horizontal transmission limited<sup>26</sup> (as compared to pathogens) and the pattern of antibiotic production correlating with resistance. Exploiting uncultured bacteria is likely to revive the Waksman platform of natural product drug discovery<sup>7</sup>. Teixobactin is a promising therapeutic candidate; it is effective against drug-resistant pathogens in a number of animal models of infection. Binding of teixobactin to WTA precursor contributes to efficient lysis and killing, due to digestion of the cell wall by liberated autolysins. This is akin to the action of another natural product with excellent killing ability, acyldepsipeptide, which converts the ClpP protease into a non-specific hydrolase that digests the cell<sup>27</sup>. These examples show that natural products evolved to exploit the inherent weaknesses of bacteria<sup>28</sup>, and additional compounds that subvert important enzymes into killing devices are likely to be discovered. Teixobactin binds to multiple targets, none of which is a protein (Fig. 3e and Extended Data Fig. 7). Polyprenyl-coupled cell envelope precursors, such as lipid II, are readily accessible on the outside of Gram-positive bacteria and represent an 'Achilles heel' for antibiotic attack<sup>28</sup>. The target of teixobactin, the pyrophosphate-sugar moiety of these molecules, is highly conserved among eubacteria. The producer is a Gram-negative bacterium, and its outer membrane will protect it from re-entry of the compound (Fig. 3e and Extended Data Fig. 7). This suggests that the producer does not employ an alternative pathway for cell wall synthesis that would protect it from teixobactin, and which other bacteria could borrow. Resistance could eventually emerge from horizontal transmission of a resistance mechanism from some soil bacterium, and given the highly conserved teixobactin binding motif, this would likely take the form of an antibiotic modifying enzyme. However, although determinants coding for enzymes attacking frequently found antibiotics such as  $\beta$ -lactams or aminoglycosides are common, they are unknown for the rare vancomycin. The recently discovered teixobactin is even less common than vancomycin. After its introduction into the clinic, it took 30 years for vancomycin resistance to appear<sup>29</sup>. The lipid II modification pathway resulting in vancomycin resistance probably originated in the producer of vancomycin, *Amycolatopsis orientalis*<sup>19</sup>. It will probably take even longer for resistance to the better-protected teixobactin to emerge. Teixobactin is the first member of a new class of lipid II binding antibiotics, structurally distinct from glycopeptides, lantibiotics<sup>30,31</sup>, and defensins<sup>32</sup>. The properties of teixobactin suggest that it evolved to minimize resistance development by target microorganisms. It is likely that additional natural compounds with similarly low susceptibility to resistance are present in nature and are waiting to be discovered.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 29 July; accepted 19 November 2014.**

**Published online 7 January; corrected online 21 January 2015 (see full-text HTML version for details).**

1. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
2. Kardos, N. & Demain, A. L. Penicillin: the medicine with the greatest impact on therapeutic outcomes. *Appl. Microbiol. Biotechnol.* **92**, 677–687 (2011).
3. Schatz, A., Bugie, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc. Soc. Exp. Biol. Med.* **55**, 66–69 (1944).
4. Spellberg, B. & Shlaes, D. Prioritized current unmet needs for antibacterial therapies. *Clin. Pharmacol. Ther.* **96**, 151–153 (2014).

5. Bush, K. *et al.* Tackling antibiotic resistance. *Nature Rev. Microbiol.* **9**, 894–896 (2011).
6. Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature Rev. Drug Discov.* **6**, 29–40 (2007).
7. Lewis, K. Antibiotics: Recover the lost art of drug discovery. *Nature* **485**, 439–440 (2012).
8. Lewis, K. Platforms for antibiotic discovery. *Nature Rev. Drug Discov.* **12**, 371–387 (2013).
9. Kaeblerlein, T., Lewis, K. & Epstein, S. S. Isolating “uncultivable” microorganisms in pure culture in a simulated natural environment. *Science* **296**, 1127–1129 (2002).
10. Nichols, D. *et al.* Use of ichip for high-throughput *in situ* cultivation of “uncultivable” microbial species. *Appl. Environ. Microbiol.* **76**, 2445–2450 (2010).
11. D’Onofrio, A. *et al.* Siderophores from neighboring organisms promote the growth of uncultured bacteria. *Chem. Biol.* **17**, 254–264 (2010).
12. Gavrish, E. *et al.* Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. *Chem. Biol.* **21**, 509–518 (2014).
13. Wilson, M. C. *et al.* An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**, 58–62 (2014).
14. Nichols, D. *et al.* Short peptide induces an “uncultivable” microorganism to grow *in vitro*. *Appl. Environ. Microbiol.* **74**, 4889–4897 (2008).
15. Sakoulas, G. *et al.* Relationship of MIC and bactericidal activity to efficacy of vancomycin for treatment of methicillin-resistant *Staphylococcus aureus* bacteremia. *J. Clin. Microbiol.* **42**, 2398–2402 (2004).
16. Kollef, M. H. Limitations of vancomycin in the management of resistant staphylococcal infections. *Clin. Infect. Dis.* **45** (Suppl 3), S191–S195 (2007).
17. Arthur, M., Depardieu, F., Reynolds, P. & Courvalin, P. Quantitative analysis of the metabolism of soluble cytoplasmic peptidoglycan precursors of glycopeptide-resistant enterococci. *Mol. Microbiol.* **21**, 33–44 (1996).
18. Bugg, T. D. *et al.* Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry* **30**, 10408–10415 (1991).
19. Marshall, C. G., Broadhead, G., Leskiw, B. K. & Wright, G. D. D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proc. Natl Acad. Sci. USA* **94**, 6480–6483 (1997).
20. D’Elia, M. A. *et al.* Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. *J. Bacteriol.* **188**, 4183–4189 (2006).
21. Bierbaum, G. & Sahl, H. G. Induction of autolysis of staphylococci by the basic peptide antibiotics Pep 5 and nisin and their influence on the activity of autolytic enzymes. *Arch. Microbiol.* **141**, 249–254 (1985).
22. O’Riordan, K. & Lee, J. C. *Staphylococcus aureus* capsular polysaccharides. *Clin. Microbiol. Rev.* **17**, 218–234 (2004).
23. Xayarath, B. & Yother, J. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J. Bacteriol.* **189**, 3369–3381 (2007).
24. Degen, D. *et al.* Transcription inhibition by the depsipeptide antibiotic salinamide A. *eLife* **3**, e02451 (2014).
25. Doroghazi, J. R. *et al.* A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nature Chem. Biol.* **10**, 963–968 (2014).
26. Forsberg, K. J. *et al.* Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**, 612–616 (2014).
27. Conlon, B. P. *et al.* Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature* **503**, 365–370 (2013).
28. Schneider, T. & Sahl, H. G. An oldie but a goodie—cell wall biosynthesis as antibiotic target pathway. *Int. J. Med. Microbiol.* **300**, 161–169 (2010).
29. Leclercq, R., Derlot, E., Duval, J. & Courvalin, P. Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N. Engl. J. Med.* **319**, 157–161 (1988).
30. Wiedemann, I. *et al.* Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *J. Biol. Chem.* **276**, 1772–1779 (2001).
31. Hasper, H. E. *et al.* An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. *Science* **313**, 1636–1637 (2006).
32. Schneider, T. *et al.* Plectasin, a fungal defensin, targets the bacterial cell wall precursor Lipid II. *Science* **328**, 1168–1172 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by NIH grant T-R01 AI085585 to K.L., by NIH grant AI085612 to A.L.S., by the Charles A. King Trust to B.P.C., and by the German Research Foundation (DFG; SCHN1284/1-2) and the German Center for Infection Research (DZIF) to T.S. and I.E. The NRS strains were provided by the Network on Antimicrobial Resistance in *Staphylococcus aureus* for distribution by BEI Resources, NIAID, NIH. Preclinical Services offered by NIAID are gratefully acknowledged. We thank H. G. Sahl for reading the manuscript and making comments, A. Makriyannis for suggestions, P. Muller, B. Berdy and S. Kaluziak for taxonomy analysis, and M. Josten for performing mass spectrometry analysis.

**Author Contributions** K.L. and T.S. designed the study, analysed results, and wrote the paper. L.L.L. designed the study and analysed results. A.J.P. designed the study, performed compound isolation and structure determination and analysed data. B.P.C. designed the study, performed susceptibility experiments and wrote the paper. D.E.H. oversaw preclinical work including designing studies and analysing data. S.E. designed cultivation experiments and analysed data. M.J., L.L. and V.A.S. designed and performed experiments on structure determination and analysed data. I.E. and A.M. designed and performed experiments on mechanism of action. A.L.S., D.R.C., C.R.F., K.A.F., W.P.M., A.G.N., A.M.Z. and C.C. performed experiments on compound production, isolation, susceptibility testing and data analysis. T.F.S. identified the biosynthetic cluster.

**Author Information** The biosynthetic gene cluster for teixobactin has been deposited with GenBank under accession number KP006601. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.L. ([k.lewis@neu.edu](mailto:k.lewis@neu.edu)).

## METHODS

**Isolation and cultivation of producing strains.** A sample of 1 g of soil sample collected from a grassy field in Maine was agitated vigorously in 10 ml of deionized H<sub>2</sub>O for 10 min. After letting the soil particulates settle for 10 min, the supernatant was diluted in molten SMS media (0.125 g casein, 0.1 g potato starch, 1 g casamino acids, 20 g bacto-agar in 1 litre of water) to achieve an average concentration of one cell per 20 µl of medium. Then 20 µl aliquots were then dispensed into the wells of an iChip. The iChip was placed in direct contact with the soil. After one month of incubation, the iChips were disassembled and individual colonies were streaked onto SMS agar to test for the ability to propagate outside the iChip and for colony purification.

**Extract preparation and screening for activity.** Isolates that grew well outside the iChip were cultured in seed broth (15 g glucose, 10 g malt extract, 10 g soluble starch, 2.5 g yeast extract, 5 g casamino acids, and 0.2 g CaCl<sub>2</sub>·2H<sub>2</sub>O per 1 litre of deionized H<sub>2</sub>O, pH 7.0) to increase biomass, followed by 1:20 dilution into 4 different fermentation broths. After 11 days of agitation at 29 °C, the fermentations were dried and resuspended in an equal volume of 100% DMSO. Then 5 µl of extracts were spotted onto a lawn of growing *S. aureus* NCTC8325-4 cells in Mueller-Hinton agar (MHA) plates. After 20 h of incubation at 37 °C, visible clearing zones indicated antibacterial activity. The extract from this isolate, which was provisionally named *Eleftheria terrae* sp., produced a large clearing zone. Although *E. terrae* sp. produced antibacterial activity under several growth conditions, the best activity (that is, largest clearing zone) was seen with R4 fermentation broth (10 g glucose, 1 g yeast extract, 0.1 g casamino acids, 3 g proline, 10 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 4 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.2 g K<sub>2</sub>SO<sub>4</sub>, 5.6 g TES free acid (2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid) per 1 litre of deionized H<sub>2</sub>O, pH 7).

**Sequencing of the strain.** Genomic DNA of *E. terrae* was isolated. Sequencing was performed at the Tufts University Core Facility (Boston, MA). A paired-end library with an insert size of approximately 800 bases was generated and sequenced using Illumina technology. The read length was 251 bases per read.

**Strain identification.** A suspension of cells was disrupted by vigorous agitation with glass beads (106 nm or smaller) and the supernatant used as template to amplify the 16S rRNA gene, using GoTaq Green Master Mix (Promega M7122), and the universal primers E8F and U1510R<sup>33</sup>. The thermocycler parameters included 30 cycles of 95 °C for 30 s, 45 °C for 30 s and 72 °C for 105 s. The amplified DNA fragment was sequenced by Macrogen USA (Cambridge, MA), and the sequence compared by BLAST to cultured isolates in the Ribosomal Database Project.

The assembled genome for *E. terrae* was submitted to the RAST genome annotation server at (<http://rast.nmpdr.org/>)<sup>34</sup> which produced a list of closest relatives with published genomes. These are *Alicyclophilus denitrificans*, *Leptothrix cholodnii*, *Methylobium petroleiphilum*, and *Rubrivivax gelatinosus*, and their genomes were downloaded from the NCBI ftp site ([ftp://ftp.ncbi.nih.gov/genomes/ASSEMBLY\\_BACTERIA/](ftp://ftp.ncbi.nih.gov/genomes/ASSEMBLY_BACTERIA/)). DNA–DNA hybridization (DDH) values of these genomes to *E. terrae* were then predicted by the Genome-to-Genome Distance calculator 2.0, formula 2, (<http://ggdc.dsmz.de/>)<sup>35–37</sup>. Note that *M. petroleiphilum* and *R. gelatinosus* are present on the phylogeny tree of *E. terrae* (Extended Data Fig. 2).

**Biosynthetic gene cluster identification.** By screening the draft genome of *E. terrae*, obtained by Illumina sequencing, many gene fragments putatively belonging to NRPS coding genes were identified. The assembly was manually edited and gap closure PCRs were performed. Sanger sequencing of the resulting fragments allowed the closure of the gene locus corresponding to the teixobactin biosynthetic gene cluster. The specificity of the adenylation domains was determined using the on-line tool NRPSpredictor2 (ref. 38).

**Strain fermentation and purification of teixobactin.** Homogenized colonies were first grown with agitation in seed broth. After 4 days at 28 °C, the culture was diluted 5% (v/v) into the R4 fermentation media, and production monitored with analytical HPLC. For scale-up isolation and purification of teixobactin, 40 litres of cells were grown in a Sartorius Biostat Cultibag STR 50/200 Bioreactor for about 7 days. The culture was centrifuged and the pellet extracted with 10 litres of 50% aqueous acetonitrile and the suspension again centrifuged for 30 min. The acetonitrile was removed from the supernatant by rotary evaporation under reduced pressure until only water remained. The mixture was then extracted twice with 5 litres of n-BuOH. The organic layer was transferred to a round bottom flask and the n-BuOH removed by rotary evaporation under reduced pressure. The resulting yellow solid was dissolved in DMSO and subjected to preparatory HPLC (SP: C18, MP: H<sub>2</sub>O/MeCN/0.1% TFA). The fractions containing teixobactin were then pooled and the acetonitrile removed by rotary evaporation under reduced pressure. The remaining aqueous mixture was then lyophilized to leave a white powder (trifluoroacetate salt). Teixobactin was then converted to a hydrochloride salt, and endotoxin removed as follows. 100 mg of teixobactin (TFA salt) was dissolved in 100 ml of H<sub>2</sub>O and 5 g of Dowex (1 × 4 Cl<sup>−</sup> form) was added and the mixture incubated for 20 min with occasional shaking. A 10 g Dowex (1 × 4 Cl<sup>−</sup> form) column was prepared and the mixture was then poured onto the prepared column

and the solution was allowed to elute slowly. This solution was then poured over a fresh 10 g Dowex (1 × 4 Cl<sup>−</sup> form) column and the resulting solution filtered through a Pall 3K Molecular Weight Centrifugal filter. The clear solution was then lyophilized to leave a white powder.

**Minimum inhibitory concentration (MIC).** MIC was determined by broth microdilution according to CLSI guidelines. The test medium for most species was cation-adjusted Mueller-Hinton broth (MHB). The same test medium was supplemented with 3% lysed horse blood (Cleveland Scientific, Bath, OH) for growing *Streptococci*. *Haemophilus* Test Medium was used for *H. influenzae* (Teknova, Hollister, CA), Middlebrook 7H9 broth (Difco) was used for mycobacteria, Schaedler-anaerobe broth (Oxoid) was used for *C. difficile*, and fetal bovine serum (ATCC) was added to MHB (1:10) to test the effect of serum. All test media were supplemented with 0.002% polysorbate 80 to prevent drug binding to plastic surfaces<sup>39</sup>, and cell concentration was adjusted to approximately 5 × 10<sup>5</sup> cells per ml. After 20 h of incubation at 37 °C (2 days for *M. smegmatis*, and 7 days for *M. tuberculosis*), the MIC was defined as the lowest concentration of antibiotic with no visible growth. Expanded panel antibacterial spectrum of teixobactin was tested at Micromyx, Kalamazoo, MI, in broth assays. Experiments were performed with biological replicates.

**Minimum bactericidal concentration (MBC).** *S. aureus* NCTC8325-4 cells from the wells from an MIC microbroth plate that had been incubated for 20 h at 37 °C were pelleted. An aliquot of the initial inoculum for the MIC plate was similarly processed. The cells were resuspended in fresh media, plated onto MHA, and the colonies enumerated after incubating for 24 h at 37 °C. The MBC is defined as the first drug dilution which resulted in a 99.9% decrease from the initial bacterial titre of the starting inoculum, and was determined to be 2 × MIC for teixobactin. Experiments were performed with biological replicates.

**Time-dependent killing.** An overnight culture of cells (*S. aureus* HG003; vancomycin intermediate *S. aureus* SA1287) was diluted 1:10,000 in MHB and incubated at 37 °C with aeration at 225 r.p.m. for 2 h (early exponential) or 5 h (late exponential). Bacteria were then challenged with antibiotics at 10 × MIC (a desirable concentration at the site of infection), oxacillin (1.5 µg ml<sup>−1</sup>), vancomycin (10 µg ml<sup>−1</sup>) or teixobactin (3 µg ml<sup>−1</sup>) in culture tubes at 37 °C and 225 r.p.m. At intervals, 100 µl aliquots were removed, centrifuged at 10,000g for 1 min and resuspended in 100 µl of sterile phosphate buffered saline (PBS). Tenfold serially diluted suspensions were plated on MHA plates and incubated at 37 °C overnight. Colonies were counted and c.f.u. per ml was calculated. For analysis of lysis, 12.5 ml of culture at A<sub>600 nm</sub> (OD<sub>600</sub>) of 1.0 was treated with 10 × MIC of antibiotics for 24 h, after which, 2 ml of each culture was added to glass test tubes and photographed. Experiments were performed with biological replicates.

**Resistance studies.** For single step resistance, *S. aureus* NCTC8325-4 at 10<sup>10</sup> c.f.u. were plated onto MHA containing 2 ×, 4 ×, and 10 × MIC of teixobactin<sup>40</sup>. After 48 h of incubation at 37 °C, no resistant colonies were detected, giving the calculated frequency of resistance to teixobactin of < 10<sup>−10</sup>. For *M. tuberculosis*, cells were cultured in 7H9 medium and plated at 10<sup>9</sup> cells per ml on 10 plates and incubated for 3 weeks at 37 °C for colony counts. No colonies were detected.

For resistance development by sequential passaging<sup>40,41</sup>, *S. aureus* ATCC 29213 cells at exponential phase were diluted to an A<sub>600 nm</sub> (OD<sub>600</sub>) of 0.01 in 1 ml of MHB supplemented with 0.002% polysorbate 80 containing teixobactin or ofloxacin. Cells were incubated at 37 °C with agitation, and passaged at 24 h intervals in the presence of teixobactin or ofloxacin at subinhibitory concentration (see Supplementary Discussion for details). The MIC was determined by broth microdilution. Experiments were performed with biological replicates.

**Mammalian cytotoxicity.** The CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) was used to determine the cytotoxicity of teixobactin. Exponentially growing NIH/3T3 mouse embryonic fibroblast (ATCC CRL-1658, in Dulbecco's Modified Eagle's medium supplemented with 10% bovine calf serum), and HepG2 cells (ATCC HB-8065, in Dulbecco's Modified Eagle's medium supplemented with 10% fetal calf serum) were seeded into a 96-well flat bottom plate, and incubated at 37 °C. After 24 h, the medium was replaced with fresh medium containing test compounds (0.5 µl of a twofold serial dilution in DMSO to 99.5 µl of media). After 48 h of incubation at 37 °C, reporter solution was added to the cells and after 2 h, the A<sub>490 nm</sub> (OD<sub>490</sub>) was measured using a Spectramax Plus Spectrophotometer. Experiments were performed with biological replicates.

**Haemolytic activity.** Fresh human red blood cells were washed with PBS until the upper phase was clear after centrifugation. The pellet was resuspended to an A<sub>600 nm</sub> (OD<sub>600</sub>) of 24 in PBS, and added to the wells of a 96-well U-bottom plate. Teixobactin was serially diluted twofold in water and added to the wells resulting in a final concentration ranging from 0.003 to 200 µg ml<sup>−1</sup>. After one hour at 37 °C, cells were centrifuged at 1,000g. The supernatant was diluted and A<sub>450 nm</sub> (OD<sub>450</sub>) measured using a Spectramax Plus Spectrophotometer. Experiments were performed with biological replicates.

**Macromolecular synthesis.** *S. aureus* NCTC8325-4 cells were cultured in minimal medium (0.02 M HEPES, 0.002 M MgSO<sub>4</sub>, 0.0001 M CaCl<sub>2</sub>, 0.4% succinic acid,



0.043 M NaCl<sub>2</sub>, 0.5% (NH<sub>4</sub>)<sub>2</sub> SO<sub>4</sub>) supplemented with 5% tryptic soy broth (TSB). Cells were pelleted and resuspended into fresh minimal medium supplemented with 5% TSB containing test compounds and radioactive precursors to a density of 10<sup>8</sup> cells per ml. The radioactive precursors were glucosamine hydrochloride, D-[6-<sup>3</sup>H(N)] (1 mCi ml<sup>-1</sup>), leucine, L-[3,4,5-<sup>3</sup>H(N)] (1 mCi ml<sup>-1</sup>), uridine, [5-<sup>3</sup>H] (1 mCi ml<sup>-1</sup>), or thymidine, [methyl-<sup>3</sup>H] (0.25 mCi ml<sup>-1</sup>) to measure cell wall, protein, RNA, and DNA synthesis, respectively. After 20 min of incubation at 37 °C, aliquots were removed, added to ice cold 25% trichloroacetic acid (TCA), and filtered using Multiscreen Filter plates (Millipore Cat. MSDVN6B50). The filters were washed twice with ice cold 25% TCA, twice with ice-cold water, dried and counted with scintillation fluid using Perkin Elmer MicroBeta TriLux Microplate Scintillation and Luminescence counter. Experiments were performed with biological replicates.

**Intracellular accumulation of UDP-*N*-acetyl-muramic acid pentapeptide.** Analysis of the cytoplasmic peptidoglycan nucleotide precursor pool was examined using *S. aureus* ATCC 29213 grown in 25 ml MHB. Cells were grown to an A<sub>600nm</sub> (OD<sub>600</sub>) of 0.6 and incubated with 120 µg ml<sup>-1</sup> of chloramphenicol for 15 min. Teixobactin was added at 1, 2.5 and 5 × MIC and incubated for another 60 min. Vancomycin (VAN; 10 × MIC), known to form a complex with lipid II, was used as positive control. Cells were collected and extracted with boiling water. The cell extract was then centrifuged and the supernatant lyophilized<sup>42</sup>. UDP-linked cell wall precursors were analysed by RP18-HPLC<sup>43</sup> and confirmed by MALDI-ToF<sup>44</sup> mass spectrometry. Experiments were performed with biological replicates.

**Cloning, overexpression and purification of *S. aureus* UppS and YbjG as His<sub>6</sub>-tag fusions.** *S. aureus* N315 *uppS* (SA1103) and *ybjB* (SA0415) were amplified using forward and reverse primers *uppS*\_FW-5'-TCGGAGGAAAGCATATGT TTAAGGAGC-3', *uppS*\_RV-5'-ATACTCTCGAGCTCCTCACTC-3', SA0415\_FW-5'-GCGCGGGATCCATAGATAGATAAAAAATTACATCAC-3' and SA0415\_RV-5'-GCGCGCTCGAGAACGCGTTGTCGTCGATGAT-3', respectively and cloned into a modified pET20 vector<sup>44</sup> using restriction enzymes NdeI (*uppS*) or BamHI (*ybjG*) and XhoI, to generate C-terminal His<sub>6</sub>-fusion proteins. Recombinant UppS-His<sub>6</sub> enzyme was overexpressed and purified as described for MurG<sup>32</sup>. For overexpression and purification of YbjG-His<sub>6</sub> *E. coli* BL21 (DE3) C43 cells transformed with the appropriate recombinant plasmid were grown in 2YT-medium (50 µg ml<sup>-1</sup> ampicillin) at 25 °C. At an A<sub>600nm</sub> (OD<sub>600</sub>) of 0.6, IPTG was added at a concentration of 1 mM to induce expression of the recombinant proteins. After 16 h, cells were harvested and resuspended in buffer A (25 mM Tris/HCl, pH 7.5, 150 mM NaCl, 2 mM β-mercaptoethanol, 30% glycerol, and 1 mM MgCl<sub>2</sub>). 2 mg ml<sup>-1</sup> lysozyme, 75 µg ml<sup>-1</sup> DNase and 75 µg ml<sup>-1</sup> RNase were added; cells were incubated for 1 h on ice, sonicated and the resulting suspension was centrifuged (20,000g, 30 min, 4 °C). Pelleted bacterial membranes were washed three times to remove remaining cytoplasmic content. Membrane proteins were solubilized in two successive steps with buffer A containing 17.6 mM *n*-dodecyl-β-D-maltoside (DDM). Solubilized proteins were separated from cell debris by centrifugation (20,000g, 30 min, 4 °C) and the supernatant containing recombinant proteins was mixed with Talon-agarose (Clontech) and purification was performed<sup>42</sup>. Purity was controlled by SDS-PAGE and protein concentration was determined using Bradford protein assay (Biorad).

***In vitro* peptidoglycan synthesis reactions.** *In vitro* peptidoglycan biosynthesis reactions were performed as described using purified enzymes and substrates<sup>32,45</sup>. The MurG activity assay was performed in a final volume of 30 µl containing 2.5 nmol purified lipid I, 25 nmol UDP-*N*-acetyl glucosamine (UDP-GlcNAc) in 200 mM Tris-HCl, 5.7 mM MgCl<sub>2</sub>, pH 7.5, and 0.8% Triton X-100 in the presence of 0.45 µg of purified, recombinant MurG-His<sub>6</sub> enzyme. Reaction mixtures were incubated for 60 min at 30 °C. For quantitative analysis 0.5 nmol of [<sup>14</sup>C]-UDP-GlcNAc (9.25 GBq mmol<sup>-1</sup>; ARC) was added to the reaction mixtures. The assay for synthesis of lipid II-Gly<sub>1</sub> catalysed by FemX was performed as described previously without any modifications<sup>32,45</sup>. Enzymatic activity of *S. aureus* PBP2 was determined by incubating 2 nmol [<sup>14</sup>C]-lipid II in 100 mM MES, 10 mM MgCl<sub>2</sub>, pH 5.5 in a total volume of 50 µl. The reaction was initiated by the addition of 5 µg PBP2-His<sub>6</sub> and incubated for 2.5 h at 30 °C. Monophosphorylation of C<sub>55</sub>-PP was carried out using purified *S. aureus* YbjG-His<sub>6</sub> enzyme as described previously for *E. coli* pyrophosphatase<sup>46</sup>, with modifications. 0.5 nmol [<sup>14</sup>C]-C<sub>55</sub>-PP (1.017 kBq) was incubated with 0.6 µg YbjG-His<sub>6</sub> in 20 mM Tris/HCl, pH 7.5, 60 mM NaCl, 0.8% Triton X-100 for 10 min at 30 °C.

In all *in vitro* assays teixobactin was added in molar ratios ranging from 0.25 to 8 with respect to the amount of [<sup>14</sup>C]-C<sub>55</sub>-PP, lipid I or lipid II and [<sup>14</sup>C]-lipid II, respectively. Synthesized lipid intermediates were extracted from the reaction mixtures with *n*-butanol/pyridine acetate, pH 4.2 (2:1; vol/vol) after supplementing the reaction mixture with 1 M NaCl and analysed by thin-layer chromatography (TLC). Quantification was carried out using phosphorimaging (Storm imaging system, GE Healthcare) as described<sup>32,45</sup>. Experiments were performed with biological replicates.

**Synthesis and purification of lipid intermediates.** Large scale synthesis and purification of the peptidoglycan precursors lipid I and II was performed<sup>45</sup>. Radiolabelled lipid II was synthesized using [<sup>14</sup>C]-UDP-GlcNAc (9.25 GBq mmol<sup>-1</sup>; ARC) as substrate. For synthesis of the lipid II variant with a terminal D-Lac residue, UDP-MurNAc-depsipeptide (Ala-Glu-Lys-Ala-Lac) was purified from *Lactobacillus casei* ATCC393. Briefly, *L. casei* was grown in MRS broth to an A<sub>600nm</sub> (OD<sub>600</sub>) of 0.6 and incubated with 65 µg ml<sup>-1</sup> of chloramphenicol for 15 min. Intracellular accumulation was achieved by incubation with Bacitracin (10 × MIC, 40 µg ml<sup>-1</sup>) in the presence of 1.25 mM zinc for another 60 min. For synthesis of lipid II ending D-Ala-D-Ser the UDP-MurNAc-pentapeptide (Ala-Glu-Lys-Ala-Ser) was used. The wall teichoic acid precursor lipid III (undecaprenyl-PP-GlcNAc) was prepared using purified TarO enzyme<sup>44</sup>. In short, purified recombinant TarO protein was incubated in the presence of 250 nmol C<sub>55</sub>-P, 2.5 µmol of UDP-GlcNAc in 83 mM Tris-HCl (pH 8.0), 6.7 mM MgCl<sub>2</sub>, 8.3% (v/v) dimethyl sulfoxide, and 10 mM *N*-lauroylsarcosine. The reaction was initiated by the addition of 150 µg of TarO-His<sub>6</sub> and incubated for 3 h at 30 °C. Lipid intermediates were extracted from the reaction mixtures with *n*-butanol/pyridine acetate (pH 4.2) (2:1; vol/vol), analysed by TLC and purified. C<sub>55</sub>-P and C<sub>55</sub>-PP were purchased from Larodan Fine Chemicals, Sweden. [<sup>14</sup>C]-C<sub>55</sub>-PP was synthesized using purified *S. aureus* UppS enzyme based on a protocol elaborated for *E. coli* undecaprenyl pyrophosphate synthase<sup>47</sup>. Synthesis was performed using 0.5 nmol [<sup>14</sup>C]-farnesyl pyrophosphate (ARC; 2.035 GBq mmol<sup>-1</sup>), 5 nmol isopentenyl pyrophosphate (Sigma-Aldrich) and 5 µg UppS enzyme in 100 mM HEPES, pH 7.6, 50 mM KCl, 5 mM MgCl<sub>2</sub>, and 0.1% Triton X-100. After 3 h of incubation at 30 °C radiolabelled C<sub>55</sub>-PP was extracted from the reaction mixture with BuOH and dried under vacuum. Product identity was confirmed by TLC analysis. Experiments were performed with biological replicates.

**Antagonization assays.** Antagonization of the antibiotic activity of teixobactin by potential target molecules was performed by an MIC-based setup in microtitre plates. Teixobactin (8 × MIC) was mixed with potential HPLC-purified antagonists (C<sub>55</sub>-P, farnesyl-PP [C<sub>15</sub>-PP; Sigma Aldrich], C<sub>55</sub>-PP, UDP-MurNAc-pentapeptide, UDP-GlcNAc [Sigma Aldrich], lipid I, lipid II, and lipid III) at a fixed molar ratio (fivefold molar excess) or at increasing concentrations with respect to the antibiotic, and the lowest ratio leading to complete antagonization of teixobactin activity was determined. *S. aureus* ATCC 29213 (5 × 10<sup>5</sup> c.f.u. per ml) were added and samples were examined for visible bacterial growth after overnight incubation. Vancomycin (8 × MIC) was used as a control. Experiments were performed with biological replicates.

**Complex formation of teixobactin.** Binding of teixobactin to C<sub>55</sub>-P, C<sub>55</sub>-PP, lipid I, lipid II, lipid II-D-Ala-D-Ser, lipid II-D-Ala-D-Lac and lipid III was analysed by incubating 2 nmol of each purified precursor with 2 to 4 nmoles of teixobactin in 50 mM Tris/HCl, pH 7.5, for 30 min at room temperature. Complex formation was analysed by extracting unbound precursors from the reaction mixture with *n*-butanol/pyridine acetate (pH 4.2) (2:1; vol/vol) followed by TLC analysis using chloroform/methanol/water/ammonia (88:48:10:1, v/v/v/v) as the solvent and detection of lipid-containing precursors by phosphomolybdic acid staining<sup>48</sup>. Experiments were performed with biological replicates.

**hERG inhibition testing.** Teixobactin was tested for inhibition of hERG activity using an IonWorks™ HT instrument (Molecular Devices Corporation), which performs electrophysiology measurements in a 384-well plate (PatchPlate). Chinese hamster ovary (CHO) cells stably transfected with hERG (cell-line obtained from Cytomx, UK) were prepared as a single-cell suspension in extracellular solution (Dulbecco's phosphate buffered saline with calcium and magnesium pH 7), and aliquots added to each well of the plate. The cells were positioned over a small hole at the bottom of each well by applying a vacuum beneath the plate to form an electrical seal. The resistance of each seal was measured via a common ground-electrode in the intracellular compartment and individual electrodes placed into each of the upper wells. Experiments were performed with three biological replicates.

**Cytochrome P450 inhibition.** Teixobactin and control compounds were incubated with human liver microsomes at 37 °C to determine their effect on five major human cytochromes P450s (CYP). The assay included probe substrates (midazolam for Cyp3A4, testosterone for Cyp3A4, tolbutamide for Cyp2C9, dextro-methorphan for Cyp2D6, S-mephentanyl for Cyp2C19, and phenacetin for Cyp1A2, 2 mM NADPH, 3 mM MgCl<sub>2</sub> in 50 mM potassium phosphate buffer, pH 7.4. The final microsomal concentration was 0.5 mg ml<sup>-1</sup>. NADPH was added last to start the assay. After ten minutes of incubation, the amount of probe metabolite in the supernatant was determined by LC/MS/MS using an Agilent 6410 mass spectrometer coupled with an Agilent 1200 HPLC and a CTC PAL chilled autosampler, all controlled by MassHunter software (Agilent). Experiments were performed with three biological replicates.

***In vitro* genotoxicity.** Teixobactin was tested in an *in vitro* micronucleus test that employs fluorescent cell imaging to assess cytotoxicity and quantify micronuclei. The assay was performed with CHO-K1 cells in the presence or absence of Aroclor

(to induce CYP activity)-treated rat liver S9 fraction (contains phase I and phase II metabolizing enzymes) to determine if any genotoxic metabolites are produced. No evidence of genotoxicity was observed with teixobactin up to  $125 \mu\text{g ml}^{-1}$  (the highest concentration tested) under either condition. Experiments were performed with three biological replicates.

**DNA binding.** Compounds were serially diluted and mixed with sheared salmon sperm DNA ( $6.6 \text{ mg ml}^{-1}$  final concentration). An aliquot was spotted onto a lawn of growing *S. aureus* NCTC 8325-4 cells, and the zones of growth inhibition measured after 20 h of growth at  $37^\circ\text{C}$ . A reduction in the inhibition zone size in the presence of DNA would indicate loss of antibacterial activity due to binding to the DNA. Experiments were performed with three biological replicates.

**Plasma protein binding.** Protein binding of teixobactin in rat plasma was determined using a Rapid Equilibrium Dialysis (RED) kit (Pierce) with LC-MS/MS analysis. Teixobactin ( $10 \mu\text{g ml}^{-1}$ ) and rat plasma in 5% dextrose containing 0.005% polysorbate 80 were added to one side of the single-use RED plate dialysis chamber having an 8kD MW cutoff membrane. Following four hours of dialysis the samples from both sides were processed and analysed by LC/MS/MS. The teixobactin concentration was determined, and the percentage of compound bound to protein was calculated. Teixobactin exhibited 84% plasma protein binding. Experiments were performed with three biological replicates.

**Microsomal stability.** The metabolic stability of teixobactin was measured in rat liver microsomes (Invitrogen/Life Technologies, CA) using NADPH Regeneration System (Promega) by monitoring the disappearance of the compound over an incubation period of two hours. Teixobactin ( $60 \mu\text{g ml}^{-1}$ ) or verapamil ( $5 \mu\text{M}$ ) serving as positive control were added to  $1 \text{ mg ml}^{-1}$  microsomes at  $37^\circ\text{C}$ . Aliquots were removed at 0 h, 0.5 h, 1 h and 2 h, and the reactions stopped by addition of 3 volumes of ice-cold acetonitrile. Samples were analysed by LC/MS/MS. Experiments were performed with three biological replicates.

**Animal studies.** All animal studies were carried out at Vivisource Laboratories, (Waltham, MA), and University of North Texas Health Science Center (Fort Worth, TX), and conformed to institutional animal care and use policies. Neither randomization nor blinding was deemed necessary for the animal infection models, and all animals were used. All animal studies were performed with female CD-1 mice, 6–8-weeks old.

**Pharmacokinetic analysis.** CD-1 female mice were injected intravenously with a single dose of 20 mg per kg in water and showed no adverse effects. Plasma samples were taken from 3 mice per time point (5, 15, 30 min; 1, 2, 4, 8 and 24 h post-dose). An aliquot of plasma sample or calibration sample was mixed with three volumes of methanol containing internal standard, incubated on ice for 5 min, and centrifuged. The protein-free supernatant was analysed by LC/MS/MS using an Agilent 6410 mass spectrometer coupled with an Agilent 1200 HPLC and a CTC PAL chilled autosampler, all controlled by MassHunter software (Agilent). After separation on a C18 reverse phase HPLC column (Agilent) using an acetonitrile-water gradient system, peaks were analysed by mass spectrometry using ESI ionization in MRM mode. The product  $m/z$  analysed was 134.1D, which provided a low limit of quantification of  $1 \text{ ng ml}^{-1}$ . The mean plasma concentration and the standard deviation from all 3 animals within each time point were calculated. PK parameters of test agent were calculated with a non-compartmental analysis model based on WinNonlin. The mean plasma concentrations from all 3 mice at each time point were used in the calculation.

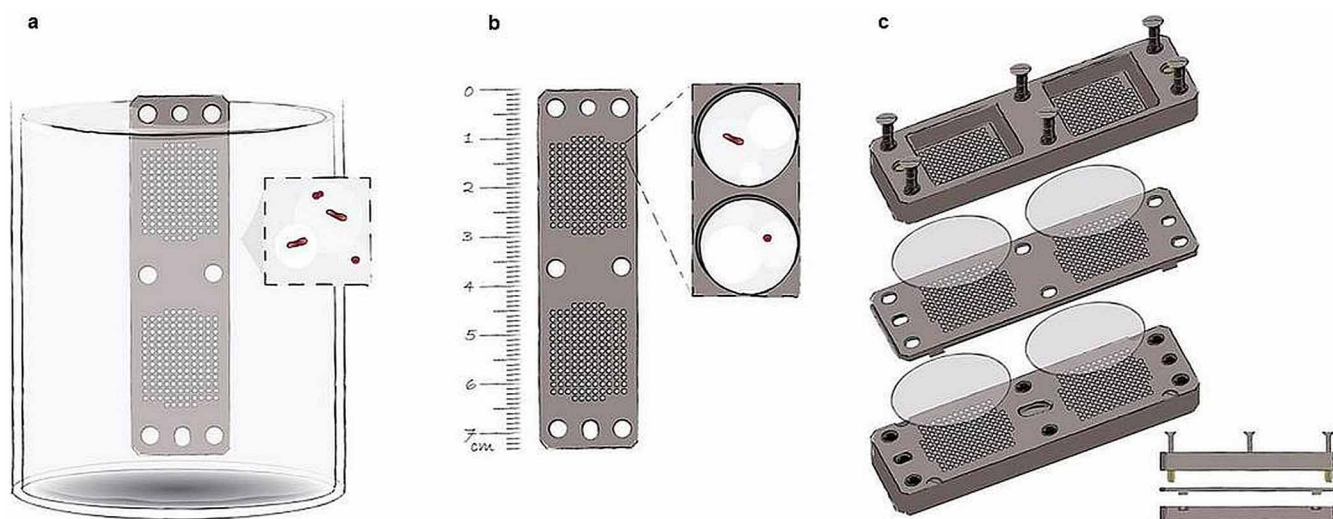
**Mouse sepsis protection model.** Teixobactin was tested against clinical isolate *S. aureus* MRSA ATCC33591 in a mouse septicemia protection assay to assess its *in vivo* bioavailability and  $\text{PD}_{50}$  (protective dose resulting in 50% survival of infected mice after 48 h). CD-1 female mice were infected with 0.5 ml of bacterial suspension ( $3.28 \times 10^7$  c.f.u. per mouse) via intraperitoneal injection, a concentration that achieves at least 90% mortality within 48 h after infection. At one hour post-infection, mice (6 per group) were treated with teixobactin at single intravenous doses of 20, 10, 5, 2.5, and 1 mg per kg. Infection control mice were dosed with vehicle or vancomycin. Survival is observed 48 h after infection and the probability determined by non-parametric log-rank test. To obtain the  $\text{PD}_{50}$ , the experiment was repeated at lower doses 5, 1, 0.5, 0.25, and 0.1 mg per kg.

**Mouse thigh infection model.** Teixobactin was tested against MRSA ATCC33591 in a neutropenic mouse thigh infection model. Female CD-1 mice were rendered

neutropenic by cyclophosphamide (two consecutive doses of 150 and 100 mg per kg delivered on 4 and 1 days before infection). Bacteria were resuspended in sterile saline, adjusted to an  $A_{625\text{nm}}$  ( $\text{OD}_{625}$ ) of 0.1, and a 0.1 ml inoculum ( $2.8 \times 10^5$  c.f.u. per mouse) injected into the right thighs of mice. At 2 h post-infection, mice received treatment with teixobactin at 1, 2.5, 5, 10 or 20 mg per kg administered in a single dose, intravenous injection (four mice per group). One group of infected mice was euthanized and thighs processed for c.f.u. to serve as the time of treatment controls. At 26 h post-infection mice were euthanized by  $\text{CO}_2$  inhalation. The right thighs were aseptically removed, weighed, homogenized, serially diluted, and plated on trypticase soy agar for c.f.u. titres.

**Mouse lung infection model.** Teixobactin was tested against *Streptococcus pneumoniae* ATCC 6301 (UNT012-2) in an immunocompetent mouse pneumonia model to determine the compound's potential to treat acute respiratory infections. CD-1 mice were infected intranasally ( $1.5 \times 10^6$  c.f.u. per mouse). The compound was delivered intravenously at 24 and 36 h post-infection, whereas amoxicillin was delivered subcutaneously at a single concentration to serve as positive control. Teixobactin was delivered at doses ranging from 0.5 to 10 mg per kg per dose (5 mice per dose). At 48 h post-infection, treated mice were euthanized, lungs aseptically removed and processed for c.f.u. counts.

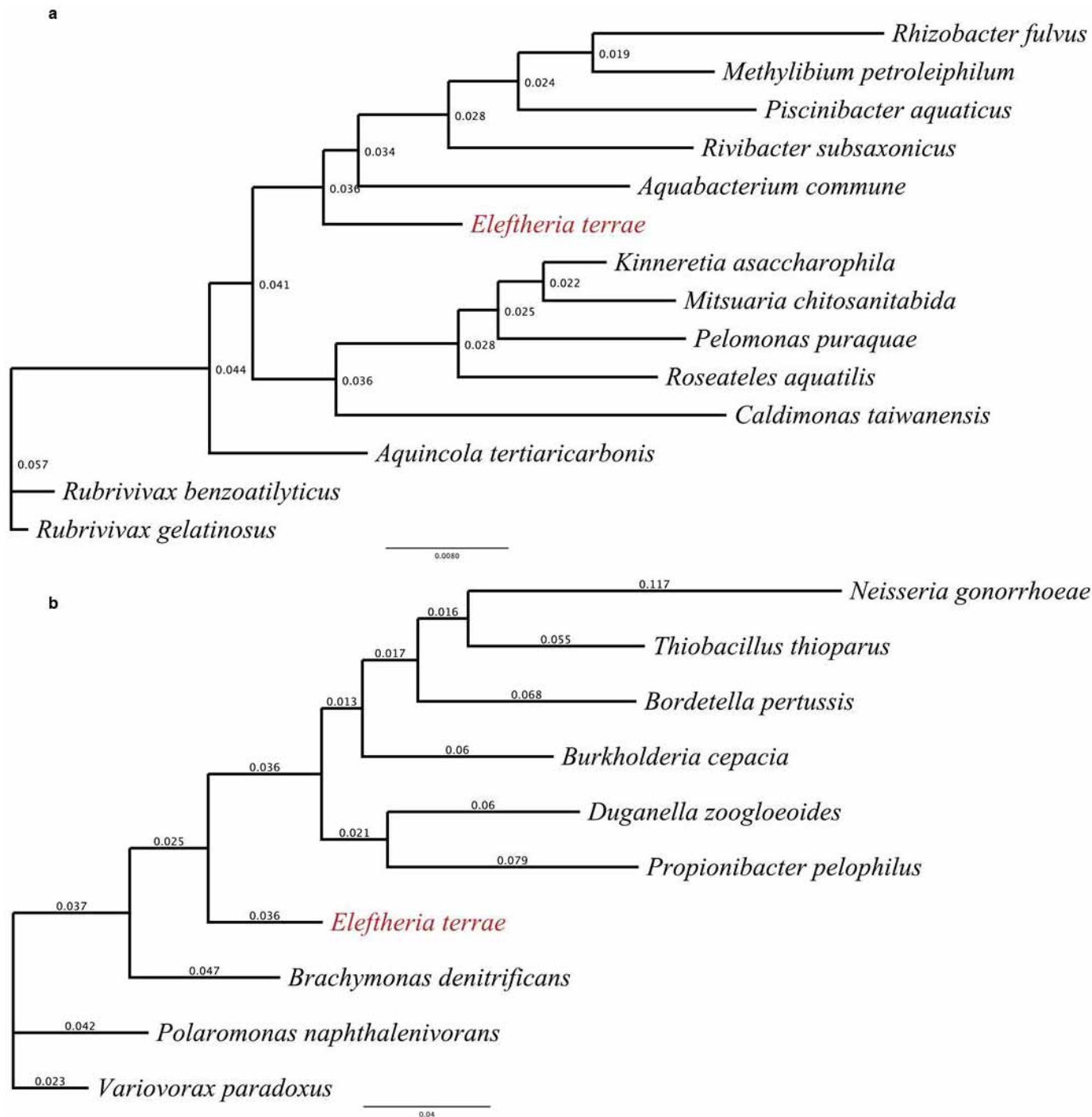
33. Baker, G. C., Smith, J. J. & Cowan, D. A. Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* **55**, 541–555 (2003).
34. Aziz, R. K. *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75 (2008).
35. Auch, A. F., Klenk, H. P. & Goker, M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand. Genomic Sci.* **2**, 142–148 (2010).
36. Auch, A. F., von Jan, M., Klenk, H. P. & Goker, M. Digital DNA–DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* **2**, 117–134 (2010).
37. Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Goker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60 (2013).
38. Röttig, M. *et al.* NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39**, W362–W367 (2011).
39. Arhin, F. F. *et al.* Effect of polysorbate 80 on oritavancin binding to plastic surfaces: implications for susceptibility testing. *Antimicrob. Agents Chemother.* **52**, 1597–1603 (2008).
40. Bogdanovich, T., Ednie, L. M., Shapiro, S. & Appelbaum, P. C. Antistaphylococcal activity of cefbiprole, a new broad-spectrum cephalosporin. *Antimicrob. Agents Chemother.* **49**, 4210–4219 (2005).
41. Metzler, K., Drlica, K. & Blondeau, J. M. Minimal inhibitory and mutant prevention concentrations of azithromycin, clarithromycin and erythromycin for clinical isolates of *Streptococcus pneumoniae*. *J. Antimicrob. Chemother.* **68**, 631–635 (2013).
42. Schneider, T. *et al.* The lipopeptide antibiotic Friulimicin B inhibits cell wall biosynthesis through complex formation with bactoprenol phosphate. *Antimicrob. Agents Chemother.* **53**, 1610–1618 (2009).
43. Brötz, H., Bierbaum, G., Reynolds, P. E. & Sahl, H. G. The lantibiotic mersacidin inhibits peptidoglycan biosynthesis at the level of transglycosylation. *Eur. J. Biochem.* **246**, 193–199 (1997).
44. Müller, A., Ulm, H., Reder-Christ, K., Sahl, H. G. & Schneider, T. Interaction of type A lantibiotics with undecaprenol-bound cell envelope precursors. *Microb. Drug Resist.* **18**, 261–270 (2012).
45. Schneider, T. *et al.* In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of *Staphylococcus aureus*. *Mol. Microbiol.* **53**, 675–685 (2004).
46. El Ghachi, M., Derbise, A., Bouhss, A. & Mengin-Lecreux, D. Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in *Escherichia coli*. *J. Biol. Chem.* **280**, 18689–18695 (2005).
47. El Ghachi, M., Bouhss, A., Blanot, D. & Mengin-Lecreux, D. The *bacA* gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. *J. Biol. Chem.* **279**, 30106–30113 (2004).
48. Sham, L. T. *et al.* Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science* **345**, 220–222 (2014).
49. Mohammadi, T. *et al.* Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J.* **30**, 1425–1432 (2011).
50. Lazarevic, V. & Karamata, D. The tagGH operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. *Mol. Microbiol.* **16**, 345–355 (1995).



**Extended Data Figure 1 | The iChip.** a–c, The iChip (a) consists of a central plate (b) which houses growing microorganisms, semi-permeable membranes on each side of the plate, which separate the plate from the environment, and two supporting side panels (c). The central plate and side panels have multiple matching through-holes. When the central plate is dipped into

suspension of cells in molten agar, the through-holes capture small volumes of this suspension, which solidify in the form of small agar plugs. Alternatively, molten agar can be dispensed into the chambers. The membranes are attached and the iChip is then placed in soil from which the sample originated.





#### Extended Data Figure 2 | 16S rRNA gene phylogeny of *Eleftheria terrae*.

**a**, The phylogenetic position of *E. terrae* within the class  $\beta$ -proteobacteria. The 16S rRNA gene sequences were downloaded from Entrez at NCBI using accession numbers retrieved from peer-reviewed publications. **b**, The phylogenetic position of *E. terrae* among its closest known relatives. The sequences were downloaded from NCBI using accession numbers retrieved from the RDP Classifier Database. For both trees, multiple sequence alignments (MSA) were constructed using ClustalW2, implementing a default Cost Matrix,

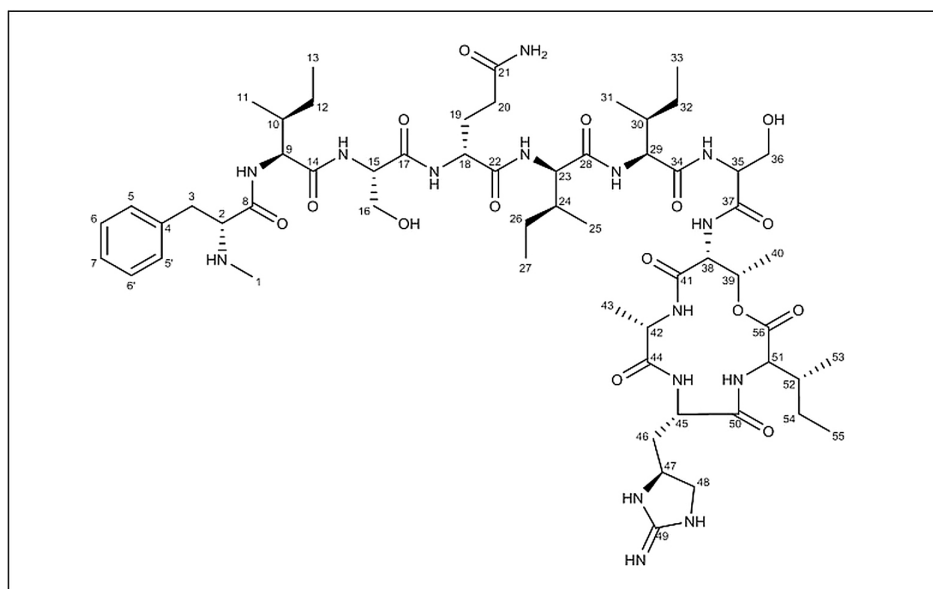
the Neighbour-Joining (NJ) clustering algorithm, as well as optimized gap penalties. Resulting alignments were manually curated and phylogenetic trees were constructed leveraging PhyML 3.0 with a TN93 substitution model and 500 Bootstrap iterations of branch support. Topology search optimization was conducted using the Subtree-Pruning-Regrafting (SPR) algorithm with an estimated Transition-Transversion ratio and gamma distribution parameters as well as fixed proportions of invariable sites.

a

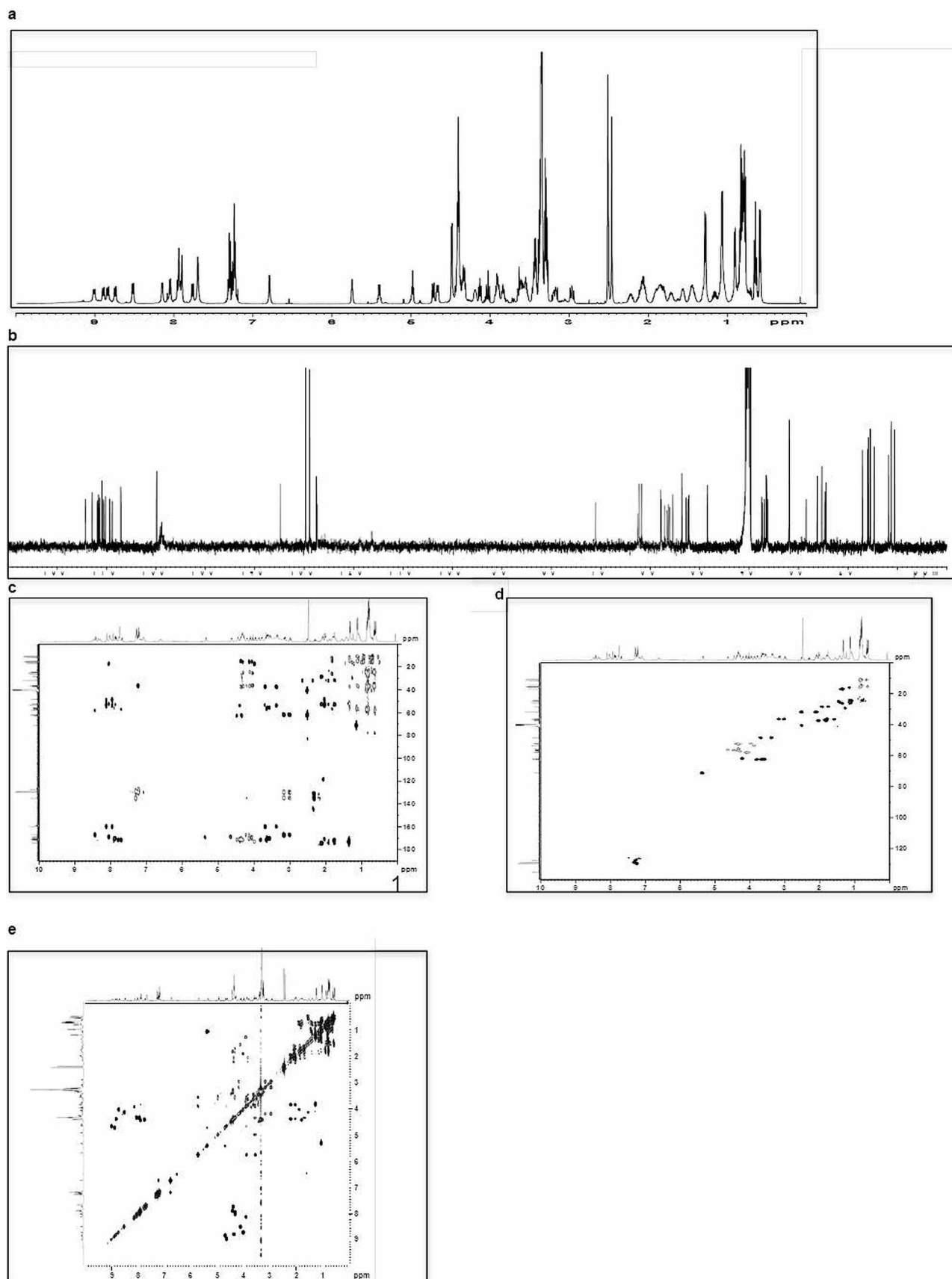
Position	$\delta_C$	$\delta_H$ (mult., J in Hz)	Position	$\delta_C$	$\delta_H$ (mult., J in Hz)
1	31.9	2.5 (3H, br s)	29	57.3	4.29 (1H, m)
2	61.9	4.21 (1H, dd, 9.4, 5.3)	29-NH		7.78 (1H, d, 8.8)
2-NH <sup>d</sup>		(2H, 9.3, 9.0, v br s)	30	36.9	1.83 (1H, m)
3	36.4	3.00 (1H, dd, 13.2, 9.4)	31	15.4 <sup>b</sup>	0.84 (3H, m)
		3.15 (1H, 13.2, 5.3)	32	25.3	1.11 (1H, m)
4	135.0				1.42 (1H, m)
5,5'	129.7	7.24 (2H, m)	33	11.2 <sup>c</sup>	0.85 (3H, m)
6,6'	128.9	7.31 (2H, m)	34	171.6 <sup>a</sup>	
7	127.5	7.27 (1H, m)	35	56.5	4.47 (1H, dt, 5.0, 5.2)
8	167.1		35-NH		8.37 (1H, d, 5.2)
9	57.9	4.12 (1H, dd, 7.8, 7.2)	36	62.7	3.64 (1H, m)
9-NH		8.43 (1H, d, 7.2)			3.80 (1H, dd, 10.8, 5.0)
10	36.5	1.56 (1H, m)	36-OH		exchanged
11	15.5	0.62 (1H, d, 6.7)	37	171.7 <sup>a</sup>	
12	24.4	0.76 (1H, m)	38	56.2	4.64 (1H, dd, 9.5, 2.2)
		1.07 (1H, m)	39	71.2	5.36 (1H, dq, 2.2, 6.4)
13	11.3	0.66 (3H, t, 7.1)	40	15.9	1.13 (3H, d, 6.4)
14	170.6		41	168.9	
15	55.6	4.34 (1H, m)	42	52.2	3.97 (1H, dq, 5.1, 7.5)
15-NH		7.88 (1H, d, 7.9)	42-NH		8.05 (1H, d, 5.1)
16	62.4	3.57 (1H, dd, 10.8, 5.6)	43	17.1	1.34 (3H, d, 7.5)
		3.63 (1H, dd, m)	44	173.1	
16-OH		exchanged	45	52.2	4.38 (1H, m)
17	170.2		45-NH		8.32 (1H, d, 9.1)
18	52.7	4.33 (1H, m)	46	37.2	2.03 (2H, m)
18-NH		7.85 (1H, d, 7.9)	47	53.5	3.90 (1H, m)
19	31.9	2.10 (2H, m)	47-NH		7.95 (1H, br s)
20	28.4	1.74 (1H, m)	48	48.3	3.36 (1H, dd, 9.4, 7.7)
		1.92 (1H, m)			3.66 (1H, t, 9.4)
21	174.4		48-NH		8.1 (1H, br s)
21-NH <sub>2</sub>		6.63 (1H, br s)	49	160.0	
		7.11 (1H, br s)	49-NH <sup>d</sup>		7.76 (2H, br s)
22	170.9 <sup>a</sup>		50	171.8 <sup>a</sup>	
23	56.8	4.36 (1H, m)	51	57.8	4.03 (1H, t, 9.4)
23-NH		7.70 (1H, d, 8.8)	51-NH		8.01 (1H, d, 9.4)
24	37.4	1.8 (2H, m)	52	36.3	1.77 (1H, m)
25	14.7 <sup>b</sup>	0.82 (3H, m)	53	16.0 <sup>b</sup>	0.81 (3H, m)
26	26.2	1.09 (1H, m)	54	24.5	0.77 (1H, m)
		1.32 (1H, m)			1.07 (1H, m)
27	10.6 <sup>c</sup>	0.82 (3H, m)	55	11.8 <sup>c</sup>	0.82 (3H, m)
28	171.4 <sup>a</sup>		56	169.3	

<sup>a</sup> Assignments may be switched due to overlap.  
<sup>b</sup> Assignments may be switched due to overlap.  
<sup>c</sup> Assignments may be switched due to overlap.  
<sup>d</sup> Appears as an ammonium salt

b

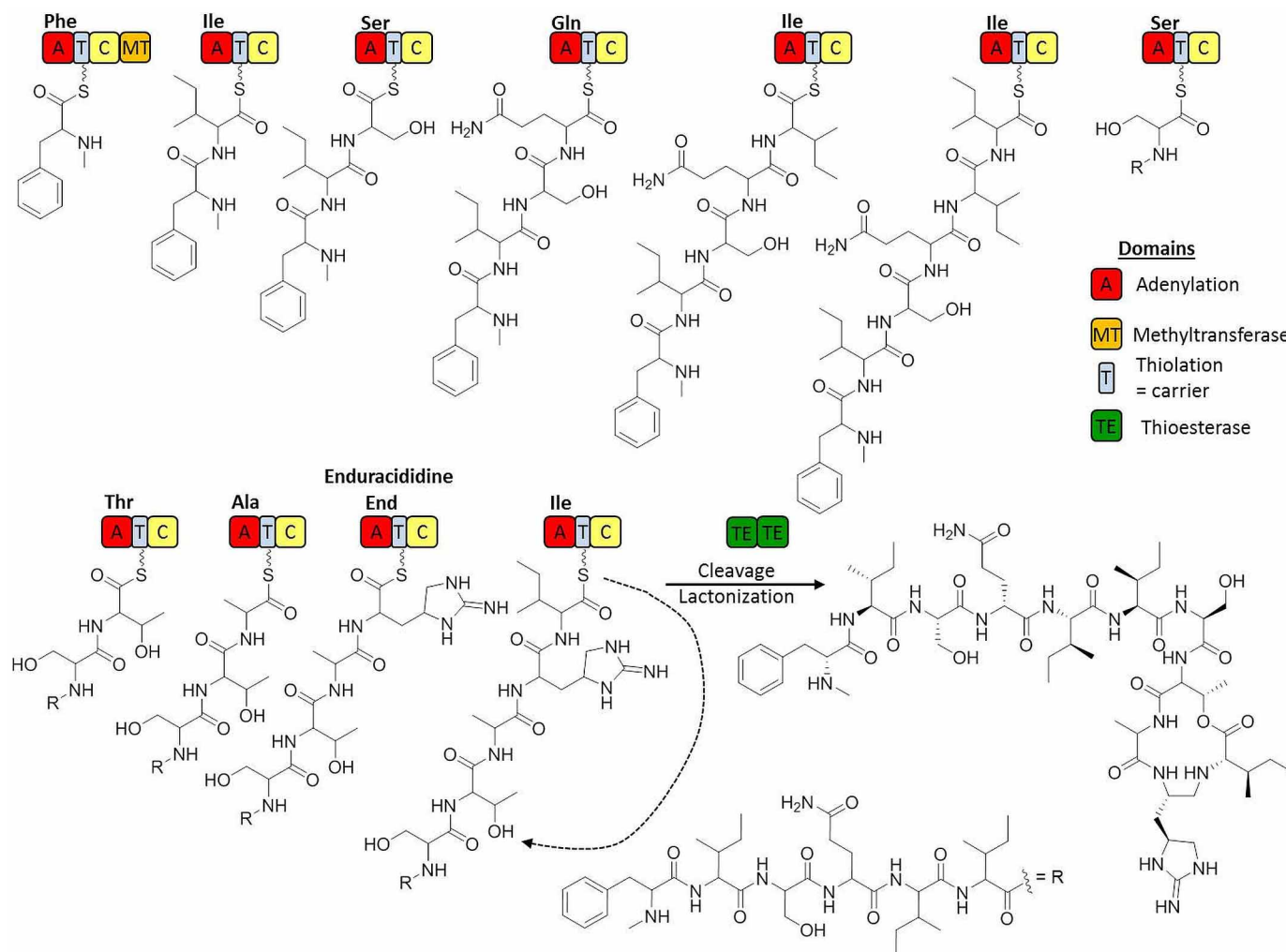


**Extended Data Figure 3 | NMR assignment of teixobactin. a,**  $^{13}\text{C}$ -NMR of teixobactin (125 MHz,  $\delta$  in p.p.m.). **b,** Structure of teixobactin with the NMR assignments.



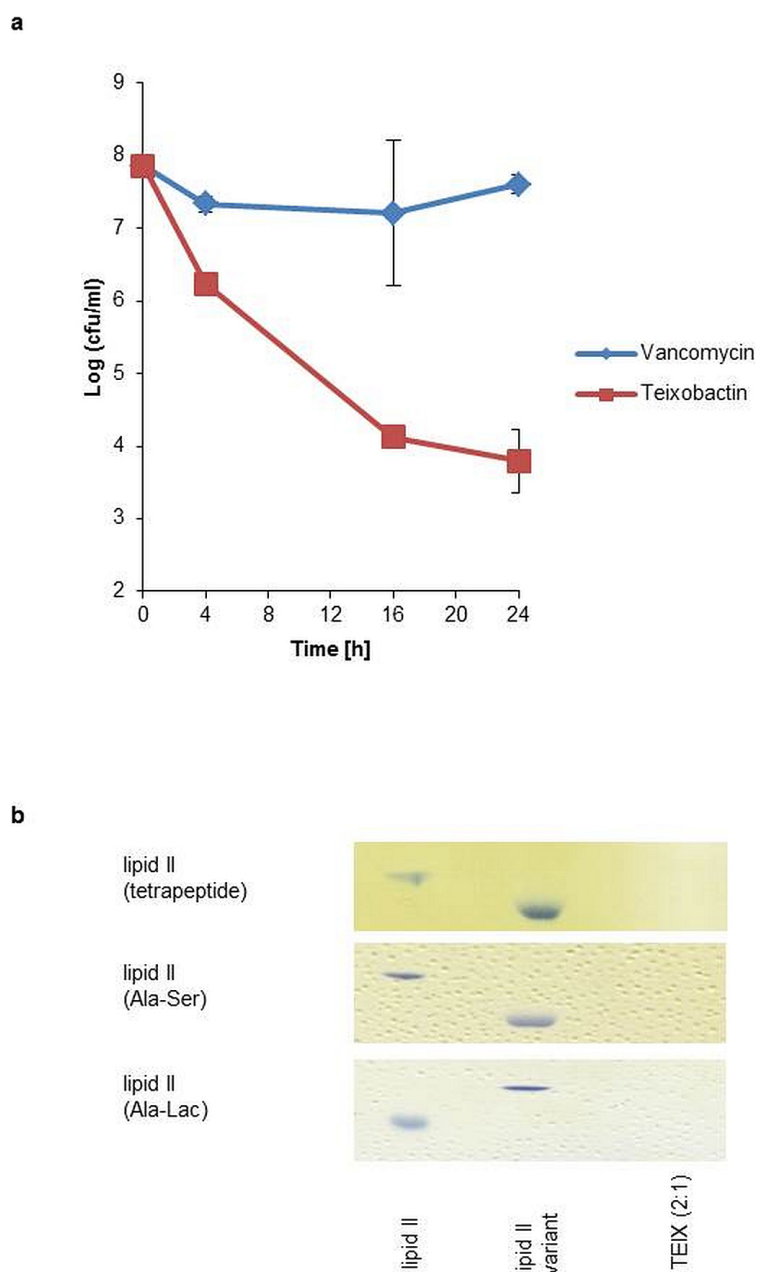
**Extended Data Figure 4** | NMR spectra of teixobactin. a,  $^{13}\text{C}$  NMR spectrum of teixobactin. b,  $^1\text{H}$  NMR spectrum. c, HMBC NMR spectrum. d, HSQC NMR spectrum. e, COSY NMR spectrum.





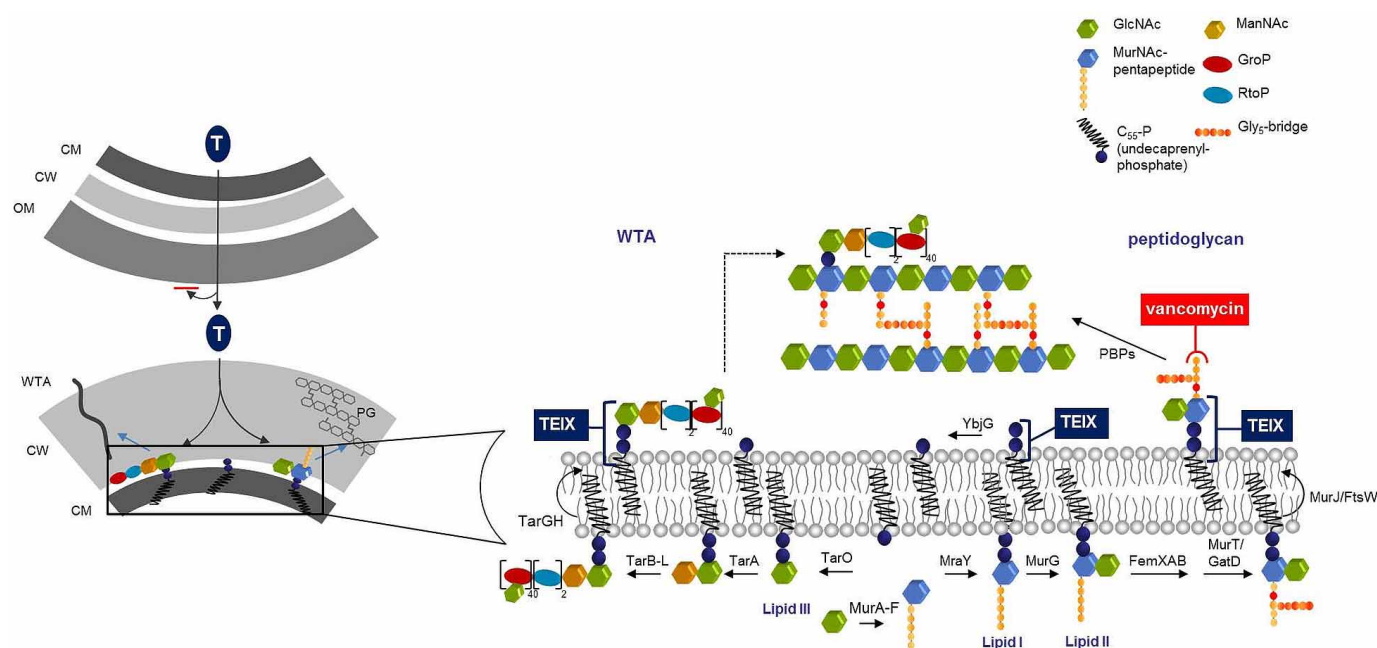
**Extended Data Figure 5 | Hypothetical biosynthesis pathway of teixobactin.** The eleven modules of the non-ribosomal peptide synthetases Txo1 and Txo2 are depicted with the growing chain attached. Each module is responsible for the incorporation of one specific amino acid in the nascent peptide chain.

The *N*-methylation of the first amino acid phenylalanine is catalysed by the methyltransferase domain in module 1. The ring closure (marked by a dashed arrow) between the last isoleucine and threonine is catalysed by the thioesterase domains during molecule off-loading, resulting in teixobactin.



**Extended Data Figure 6 | Teixobactin activity against vancomycin-resistant strains.** **a**, Vancomycin intermediate *S. aureus* (VISA) were grown to late exponential phase and challenged with vancomycin or teixobactin. Cell numbers were determined by plating for colony counts. Data are representative of 3 independent experiments  $\pm$  s.d. **b**, Complex formation of teixobactin with cell wall precursor variants as formed by vancomycin-resistant strains.

Purified lipid intermediates with altered stem peptides were incubated with teixobactin at a molar ratio of 2:1 (TEIX:lipid II variant). Reaction mixtures were extracted with BuOH/PyrAc and binding of teixobactin to lipid II variants is indicated by its absence on the thin-layer chromatogram. Migration behaviour of unmodified lipid II is used for comparison. The figure is representative of 3 independent experiments.



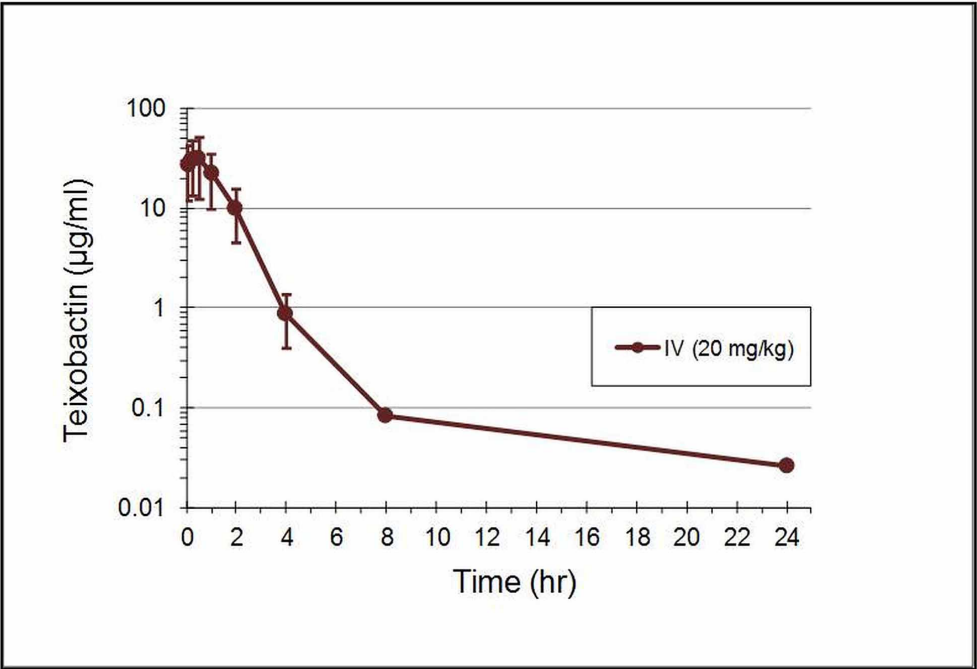
### Extended Data Figure 7 | Model for the mechanism of action of teixobactin.

Inhibition of cell wall synthesis by teixobactin. Lipid II, precursor of peptidoglycan, is synthesized in the cytoplasm and flipped to the surface of the inner membrane by MurJ<sup>48</sup> or FtsW<sup>49</sup>. Lipid III, a precursor of wall teichoic acid (WTA), is similarly formed inside the cell and WTA lipid-bound precursors are translocated across the cytoplasmic membrane by the ABC-transporter TarGH<sup>50</sup>. Teixobactin (TEIX) forms a stoichiometric complex with cell wall precursors, lipid II and lipid III. Abduction of these building blocks simultaneously interrupts peptidoglycan (right), WTA (left) biosynthesis as

well as precursor recycling. Binding to multiple targets within the cell wall pathways obstructs the formation of a functional cell envelope. Left panel, teixobactin targeting and resistance. The producer of teixobactin is a Gram-negative bacterium which is protected from this compound by exporting it outside of its outer membrane permeability barrier. The target Gram-positive organisms do not have an outer membrane. CM, cytoplasmic membrane; CW, cell wall; OM, outer membrane; LTA, lipoteichoic acid; WTA, wall teichoic acid.



a



b

PK parameter	Definition	Value
C0 (µg/mL)	Initial concentration	27.2
AUC to Last (µg-hr/mL)	Area Under Curve to last time point	57.8
t1/2 (hr)	Half life	4.7
Total CL (mL/hr)	Clearance	6.9
Total CL (mL/min/kg)	Clearance	5.8
V (mL)	Volume of Distribution	47
Vss (mL)	Volume of Distribution at steady state	9.7
MRTINF (hr)	Mean residence time	1.4
Last Time point (hr)	–	24

**Extended Data Figure 8 | Pharmacokinetic analysis of teixobactin.** **a**, The mean plasma concentrations of teixobactin after a single i.v. injection of 20 mg per kg teixobactin (3 mice per time point). Data are the mean of plasma concentration, and error bars represent the standard deviation from 3 animals in each time point. **b**, Pharmacokinetic parameters of teixobactin calculated with a non-compartmental analysis model based on WinNonlin.

Extended Data Table 1 | Antibacterial spectrum of teixobactin

Organism	MIC (µg/mL)	Organism	MIC (µg/mL)
<b><i>Staphylococcus aureus</i></b>		<b><i>Streptococcus pneumoniae</i></b>	
ATCC 29213 (MSSA)	0.16-0.31	ATCC BAA 255	0.05
NCTC 8325 (MSSA)	0.08-0.31	VL-172	0.15
ATCC 33591 (MRSA)	0.16-0.31	VL-190	0.15
NRS54 (MRSA)	0.078-0.16	ATCC 10813	0.08
NRS108 (MRSA, also synergic <sup>R</sup> )	0.16	ATCC 6303	0.02-0.04
NRS269 (MRSA, also tigecycline <sup>R</sup> )	0.16-0.31	BAA 1407	0.04
ATCC 700699 (GISA)	0.31	<b><i>Bacillus anthracis</i></b>	
<b><i>S. epidermidis</i></b>		Sterne	0.02
ATCC 35984=NRS101 ( <i>mecA</i> positive)	0.078-0.16	<b><i>B. anthracis</i> BB Resources<sup>a</sup> Isolates</b>	
NRS8 ( <i>mecA</i> positive)	0.16-0.31	NR-36 NRS 1008	≤0.06
NRS34 ( <i>mecA</i> positive)	0.16	NR-38 Pasteur Vaccine No.1	≤0.06
<b><i>S. haemolyticus</i></b>		NR-41 Graves	≤0.06
NRS9 ( <i>mecA</i> positive)	0.08	NR-46 46-PY-5	≤0.06
NRS69 ( <i>mecA</i> positive)	0.16	NR-411 Ames, A0462	≤0.06
<b><i>Enterococcus</i></b>		NR-412 Kruger B	0.06-0.125
<i>E. faecium</i> BM4147 ( <i>aac</i> (6')-le- <i>aph</i> (2''), <i>van</i> <sup>R</sup> )	0.31	NR-413 CNEVA	0.125
<i>E. faecium</i> E4sol ( <i>vancomycin</i> <sup>S</sup> )	0.31	NR-414 Vollum	≤0.06
<i>E. faecalis</i> ATCC 51575 ( <i>vancomycin</i> <sup>R</sup> )	0.31-0.63	NR-415 WNA	≤0.06
<i>E. faecalis</i> M192 ( <i>vancomycin</i> <sup>S</sup> )	0.63	NR-3838 Ames	≤0.06
<b><i>Mycobacterium</i></b>		<b>Other Gram-positive</b>	
<i>Mycobacterium smegmatis</i> mc <sup>2</sup> 155	0.31	<i>Streptococcus pyogenes</i> ATCC19615	0.31
<i>M. tuberculosis</i> H37Rv	0.125	<i>S. warneri</i> NRS138	0.02
<i>M. tuberculosis</i> (clinical isolate 70)	0.125	<i>Bacillus subtilis</i> 1A1	0.02
<i>M. tuberculosis</i> (clinical isolate 76)	0.125-0.25	<i>Clostridium difficile</i> CD196	0.005
<i>M. tuberculosis</i> (clinical isolate 82)	0.125-0.25	<i>Propionibacterium acnes</i> ATCC6919	0.078
<i>M. tuberculosis</i> (clinical isolate 102)	0.25		
<b>Gram-negative</b>			
<i>Haemophilus influenzae</i> SJ7	2.5	<i>E. coli</i> K12	25
<i>Klebsiella pneumoniae</i> ATCC 700603	20	<i>E. coli</i> W0153 (AB1157; <i>asmB1</i> $\Delta$ <i>tolC</i> :: <i>kan</i> )	2.5
<i>Pseudomonas aeruginosa</i> PA-01	>100	<i>E. coli</i> W0159 (AB1157; <i>asmB1</i> $\Delta$ <i>rfaC</i> :: <i>kan</i> )	2.5
<i>Klebsiella pneumoniae</i> ATCC 43816	>40	<i>E. coli</i> ATCC 25922	25
<i>Yersinia pestis</i> KIM 100 deletion <i>pDC1</i>	50-100	<i>E. coli</i> mutS	25
<i>Neisseria gonorrhoeae</i>	25	<i>Bacteriodes fragilis</i> ATCC 25825	200

Organism	# Isolates	Drug	MIC Range (µg/mL)	MIC <sub>90</sub> (µg/mL)	Organism	# Isolates	Drug	MIC Range (µg/mL)	MIC <sub>90</sub> (µg/mL)
<i>Staphylococcus aureus</i> MSSA	20	Teixobactin	0.06-0.25	0.25	<i>Enterococcus faecalis</i> 50% VRE	10	Teixobactin	0.5-1	0.5
		Linezolid	2-4	4			Linezolid	1-2	2
		Vancomycin	0.5-1	1			Vancomycin	0.5->32	>32
		Daptomycin	0.12-0.25	0.25			Daptomycin	0.12-8	0.5
<i>S. aureus</i> MRSA	20	Teixobactin	0.06-0.5	0.25	<i>E. faecium</i> 50% VRE	10	Teixobactin	0.25-1	1
		Linezolid	2-4	4			Linezolid	1-2	2
		Vancomycin	0.5-2	1			Vancomycin	0.5->32	>32
		Daptomycin	0.12-0.25	0.25			Daptomycin	0.5-1	1
<i>S. aureus</i> VISA	10	Teixobactin	0.12-1	0.5	<i>Streptococcus pneumoniae</i> 28.6% PSSP, 33.3% PRSP, 38.1% PRSP	20	Teixobactin	≤0.03-0.06	≤0.03
		Linezolid	1-4	2			Linezolid	0.25-1	1
		Vancomycin	1-8	8			Vancomycin	0.12-0.5	0.25
		Daptomycin	0.25-1	1			Daptomycin	≤0.03-0.06	≤0.03
<i>S. aureus</i> Daptomycin <sup>NS</sup>	5	Teixobactin	0.12-0.5	-	<i>S. pyogenes</i>	10	Teixobactin	≤0.03-0.06	0.06
		Linezolid	1-32	-			Linezolid	0.5-1	1
		Vancomycin	0.5-8	-			Vancomycin	0.25	0.25
		Daptomycin	2-8	-			Daptomycin	≤0.03-0.06	0.06
<i>S. aureus</i> Linezolid <sup>R</sup>	5	Teixobactin	0.12-0.5	-	<i>S. agalactiae</i>	10	Teixobactin	0.06-0.12	0.12
		Linezolid	16->32	-			Linezolid	0.5-2	1
		Vancomycin	1	-			Vancomycin	0.25-0.5	0.5
		Daptomycin	0.25-0.5	-			Daptomycin	0.06-0.25	0.12
<i>S. epidermidis</i>	20	Teixobactin	≤0.03-0.25	0.12	Viridans Group Streptococci <sup>1</sup>	5	Teixobactin	≤0.03-0.12	-
		Linezolid	1-8	2			Linezolid	0.5-1	-
		Vancomycin	1-2	2			Vancomycin	0.25-0.5	-
		Daptomycin	0.06-0.25	0.25			Daptomycin	0.06-0.25	-

a. Antibacterial spectrum of teixobactin. MIC was determined by broth microdilution. <sup>a</sup>*B. anthracis* BB resources isolates are from NIH Biodefense and Emerging Infections Research Resources repository.

b. Antibacterial activity of teixobactin and known drugs against contemporary clinical isolates. <sup>1</sup>In the Viridans Group Streptococci, one isolate of each of the following was tested *S. sanguis*, *S. mitis*, *S. anginosus*, *S. intermedius* and *S. salivarius*. PISP, penicillin-intermediate *S. pneumoniae*; PRSP, penicillin-resistant *S. pneumoniae*; PSSP, penicillin-sensitive *S. pneumoniae*.

Extended Data Table 2 | Antagonization of the antimicrobial activity of teixobactin by cell wall precursors

a

antagonist	C <sub>55</sub> -P	C <sub>55</sub> -PP	C <sub>15</sub> -PP	lipid I	lipid II	lipid III	UDP-MurNAc-pentapeptide	UDP-GlcNAc
teixobactin	-	+	+	+	+	+	-	-
vancomycin	-	-	nd	+	+	-	nd	nd

(+) antibiotic activity antagonized, (-) antibiotic activity unaffected, (nd) not determined

b

lipid intermediate	molar ratio of precursor to teixobactin					
	0 x	0.5 x	1 x	2.5 x	5 x	10 x
lipid II	-	+	+	+	+	+
C <sub>55</sub> -PP	-	-	-	-	+	+

**a**, *S. aureus* ATCC 29213 was incubated with teixobactin and vancomycin at 8 × MIC in nutrient broth in a microtitre plate, and growth was measured after a 24 h incubation at 37 °C. Putative HPLC-purified antagonists (undecaprenyl-phosphate [C<sub>55</sub>-P], farnesyl-pyrophosphate [C<sub>15</sub>-PP], undecaprenyl-pyrophosphate [C<sub>55</sub>-PP], UDP-MurNAc-pentapeptide, UDP-GlcNAc, lipid I, lipid II, and lipid III) were added in a fivefold molar excess with respect to the antibiotic. **b**, Teixobactin at 8 × MIC was exposed to increasing concentrations of putative antagonistic lipid intermediates. Experiments were performed with biological replicates.



# Endophilin marks and controls a clathrin-independent endocytic pathway

Emmanuel Boucrot<sup>1,2</sup>, Antonio P. A. Ferreira<sup>2</sup>, Leonardo Almeida-Souza<sup>1</sup>, Sylvain Debard<sup>2,3</sup>, Yvonne Vallis<sup>1</sup>, Gillian Howard<sup>1</sup>, Laetitia Bertot<sup>4</sup>, Nathalie Sauvonnet<sup>4</sup> & Harvey T. McMahon<sup>1</sup>

**Endocytosis is required for internalization of micronutrients and turnover of membrane components. Endophilin has been assigned as a component of clathrin-mediated endocytosis. Here we show in mammalian cells that endophilin marks and controls a fast-acting tubulovesicular endocytic pathway that is independent of AP2 and clathrin, activated upon ligand binding to cargo receptors, inhibited by inhibitors of dynamin, Rac, phosphatidylinositol-3-OH kinase, PAK1 and actin polymerization, and activated upon Cdc42 inhibition. This pathway is prominent at the leading edges of cells where phosphatidylinositol-3,4-bisphosphate—produced by the dephosphorylation of phosphatidylinositol-3,4,5-triphosphate by SHIP1 and SHIP2—recruits lamellipodin, which in turn engages endophilin. This pathway mediates the ligand-triggered uptake of several G-protein-coupled receptors such as  $\alpha_{2a}$ - and  $\beta_1$ -adrenergic, dopaminergic D3 and D4 receptors and muscarinic acetylcholine receptor 4, the receptor tyrosine kinases EGFR, HGFR, VEGFR, PDGFR, NGFR and IGF1R, as well as interleukin-2 receptor. We call this new endocytic route fast endophilin-mediated endocytosis (FEME).**

The BAR (Bin/amphiphysin/Rvs)-domain-containing protein endophilin-A (hereafter called endophilin) is an endocytic protein that recruits dynamin and synaptojanin<sup>1</sup>, and the disruption of which has a profound effect on endocytosis in mice, nematodes and flies<sup>2–5</sup>. Evidence for its involvement in clathrin-mediated endocytosis came from BAR domain and antibody injection experiments in lamprey synapses, resulting in an accumulation of early- and late-stage clathrin-coated pits and a corresponding decrease in the number of synaptic vesicles<sup>6–9</sup>. Furthermore, endophilin has been detected on clathrin-coated pits by immuno-electron microscopy and fluorescence microscopy<sup>5,9–12</sup>. However, several lines of evidence suggest a peripheral role for endophilin in clathrin-mediated endocytosis: it was only detected in one-quarter of forming clathrin-coated pits<sup>10,12</sup>; triple knockdown (TKD) of all three endophilin-A proteins (A1, A2 and A3) by RNA interference (RNAi) does not affect the uptake of the clathrin-mediated endocytosis cargo transferrin<sup>13</sup>; and the kinetics of synaptic vesicle retrieval in goldfish bipolar neurons inhibited by endophilin N-BAR domain injection (BAR domain with its amino-terminal amphipathic helix extensions) was distinct from that inhibited by a peptide inhibitor of clathrin-mediated endocytosis<sup>14,15</sup>. Interestingly, endophilin binds directly to a proline-rich motif present in the third intracellular loop (TIL) of the  $\beta_1$ -adrenergic receptor ( $\beta_1$ -AR) but not the  $\beta_2$ -AR (ref. 16), both being members of the G-protein-coupled receptor (GPCR) superfamily. Endophilin also binds to EGFR and HGFR through CIN85 and Cbl (refs 17, 18). Here we reveal that endophilin is not simply a peripheral component of clathrin-mediated endocytosis but principally marks and controls a distinct endocytic pathway.

## Endophilin-dependent endocytosis of GPCR

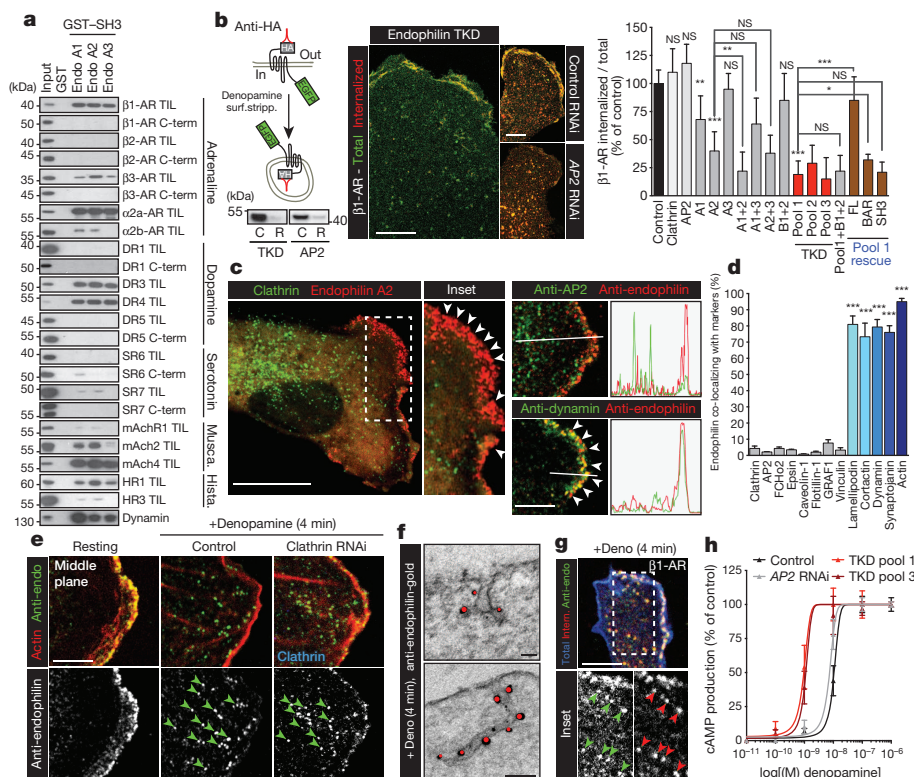
Looking for cargoes that can bind to the SH3 domain of endophilin, we found that in addition to  $\beta_1$ -AR, endophilin bound strongly to discrete proline-rich motifs within the TIL of the  $\alpha_{2a}$ -AR, dopaminergic receptors 3 and 4 (DR3 and DR4) and to the muscarinic acetylcholine receptor 4 (mAChR4), and weakly to the histaminergic receptors 1 and 3 but not

to the other members of the amine GPCR sub-family tested (Fig. 1a and Extended Data Fig. 1a–e). The internalization of  $\beta_1$ -AR in an antibody-feeding assay was strongly decreased in endophilin TKD but not in clathrin or AP2-depleted cells, suggesting an endocytic route independent of clathrin-mediated endocytosis (Fig. 1b). Among the five endophilin genes, A2 was the most important in BSC1 cells, followed by A1, with little or no role for A3, B1 and B2. The uptake of  $\beta_1$ -AR required full-length A2 and its N-BAR or SH3 domains could not rescue the endophilin TKD (Fig. 1b). Consistent with the binding data, only internalization of  $\beta_1$ -AR and  $\alpha_{2a}$ -AR, DR3, DR4, and mAChR4 was significantly reduced in endophilin TKD cells (Extended Data Fig. 1m, n). These endocytic defects were only observed after ligand stimulation, implicating endophilin in receptor-mediated endocytosis initiated upon receptor activation by their ligands. Endophilin TKD cells showed unperturbed uptake of transferrin and LDL receptors (cargo for clathrin-mediated endocytosis), but clathrin-coated pits formed at a slightly slower rate (Extended Data Fig. 1f–i), perhaps related to a reduced synaptojanin recruitment to membranes and thus an increase in phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) levels<sup>5,19</sup>. Surface  $\beta_1$ -AR accumulated mostly at the leading edge of endophilin TKD cells, which coincided with the localization of most of the endophilin signal in control cells (Fig. 1c, Extended Data Figs 1j–l and 2 and Supplementary Video 1). Consistent with the virtual absence of clathrin-coated pits and vesicles at the leading edge (Extended Data Fig. 2e), endogenous endophilin showed little co-localization with AP2 ( $2 \pm 0.1\%$  of endogenous AP2 spots), much lower than observed for endophilin overexpression data ( $\sim 5\%$  on snapshot confocal images and a brief flash just before budding in  $\sim 25\%$  of recorded clathrin-coated vesicle budding kymographs, in agreement with previous studies<sup>10,12</sup>). In confluent cells, which lack leading edges, endophilin puncta were interspersed with clathrin-coated pits on the ventral surface (Extended Data Fig. 2i, j and Supplementary Video 2).

Endophilin co-localized at the leading edge with its binding proteins (dynamin, synaptojanin, lamellipodin) and with F-actin but not with the clathrin-independent endocytic proteins caveolin-1, flotillin1/2 or

<sup>1</sup>MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK. <sup>2</sup>Institute of Structural and Molecular Biology, University College London & Birkbeck College, London WC1E 6BT, UK.

<sup>3</sup>Department of Biology, Ecole Normale Supérieure de Cachan, 94235 Cachan, France. <sup>4</sup>Institut Pasteur, Unité de Pathogénie Moléculaire Microbienne, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France.



**Figure 1 | Clathrin-independent, endophilin-dependent endocytosis of GPCR.** **a**, Pull-down experiments with EGFP-tagged third intracellular loops (TIL) or C termini of indicated GPCRs. **b**, Left: antibody feeding assay using double-tagged  $\beta_1$ -AR. HA antibodies are internalized after denopamine stimulation and non-internalized antibody was removed by surface stripping (surf.stripp.). Bottom: immunoblots from control (C) and RNAi (R) endophilin TKD and AP2-depleted cells. Right: confocal microscopy images of internalized (red) and total (green)  $\beta_1$ -AR-EGFP (images are representative of at least ten captures from three independent experiments). Bar graphs show the amounts of internalized receptors corrected for their total levels in the indicated samples (means  $\pm$  s.e.m. from three independent experiments; NS, non-significant; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; one-way ANOVA and Dunnett's test versus control or versus A2 or TKD pool 1, as indicated by the brackets). FL, full length. **c**, Confocal sections showing localization of endophilin-A2-RFP (left) or endogenous endophilin (right) with clathrin (EGFP-LCa), AP2 or dynamin in BSC1 cells (images are

representative of at least ten captures from three independent experiments). Intensity profiles were acquired along the indicated lines. Arrowheads show leading edge. **d**, Percentage of endogenous endophilin spots co-localizing with the named markers (means  $\pm$  s.e.m. from three independent experiments; \*\*\* $P < 0.001$ , one-way ANOVA and Dunnett's test versus control IgG). **e**, Confocal sections of cells immunostained for endophilin (green) and actin (red) following the indicated treatments. Arrowheads show endophilin-positive assemblies (EPAs) (images are representative of at least ten captures from three independent experiments). **f**, Immuno-electron microscopy of BSC1 cells immunostained for endophilin (red) (images are representative of ten captures from two independent experiments). **g**, Co-staining of internalized  $\beta_1$ -AR and endogenous endophilin after stimulation. Blue, total  $\beta_1$ -AR (images are representative of ten captures from two independent experiments). **h**, cAMP production following stimulation with denopamine (means  $\pm$  s.e.m. from three independent experiments). Scale bars: 20  $\mu$ m (c, left), 10  $\mu$ m (b), 5  $\mu$ m (c, right, e, g) and 250 nm (f).

GRAF1 (refs 20–22) (Fig. 1c, d and Extended Data Fig. 3). As the depletion of the latter did not affect endophilin signals or internalization of the five GPCRs identified, it suggested a hitherto novel endocytic process.

Very shortly ( $t_{1/2} = 7$  s) after addition of  $\beta_1$ -AR but not  $\beta_2$ -AR agonists, numerous endophilin-RFP-positive assemblies (EPAs) emanated from the cell edges and rapidly ( $1.62 \pm 0.29 \mu\text{m s}^{-1}$ ) travelled to the perinuclear area (Extended Data Fig. 4a, b and Supplementary Video 3). The formation of intracellular tubules and vesicles positive for endogenous endophilin after  $\beta_1$ -AR-specific activation was confirmed by confocal and electron microscopy (Fig. 1e–g and Extended Data Fig. 4c–o). EPAs were quantified on middle confocal planes and the vesicles/tubules formed were probably detached as they were no longer accessible to externally applied dyes.

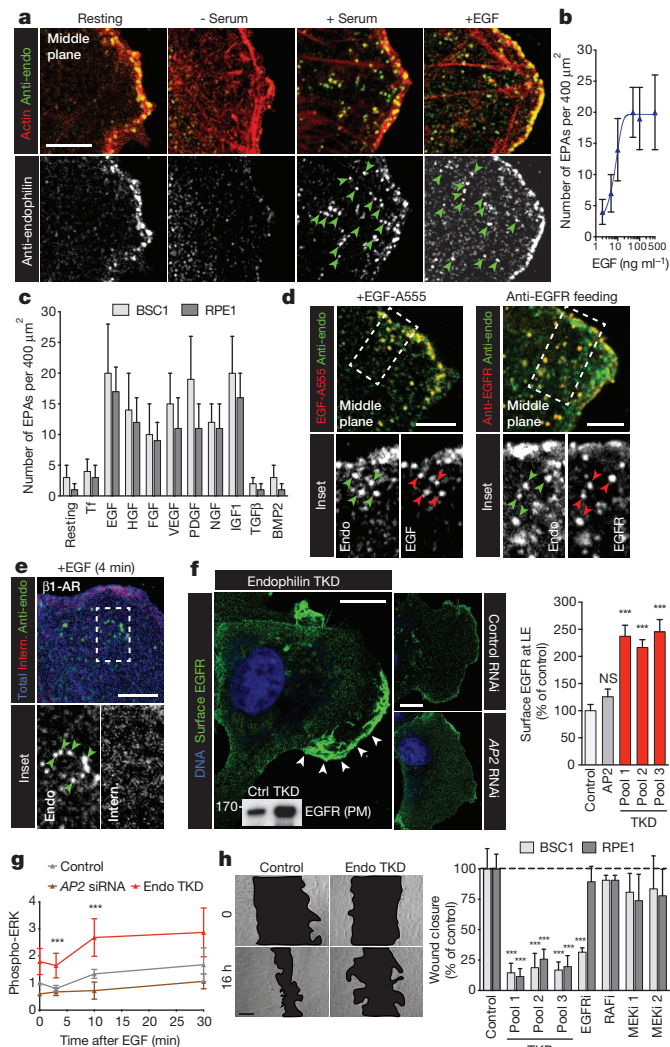
Consistently with the absence of co-localization, EPA formation was unaffected by the inhibition of clathrin-mediated endocytosis by AP180 carboxy terminus overexpression or by RNAi of the clathrin-mediated endocytosis components clathrin heavy chain or AP2 (Extended Data Fig. 4g and j). As expected from the binding data, only the overexpression of TILs from the receptors that bound to endophilin perturbed EPA formation (Extended Data Fig. 5a, b). EPAs formed upon  $\beta_1$ -AR activation were confirmed as authentic endocytic carriers because in the antibody-feeding assay they contained anti-HA antibodies (Fig. 1g).

Finally, signal transduction downstream of  $\beta_1$ -AR (cAMP production and PKA-mediated CREB phosphorylation) was elevated in endophilin TKD cells (Fig. 1h and Extended Data Fig. 5c, d), suggesting that EPAs mediate the downregulation of activated  $\beta_1$ -AR. Even though  $\beta$ -arrestin, a key clathrin-mediated endocytosis adaptor for  $\beta_2$ -AR, was located at the edge of cells, it was not detected on EPAs, nor did its overexpression or depletion perturb EPA formation after  $\beta_1$ -AR stimulation (Extended Data Fig. 5e–h). Given that these EPAs form rapidly in an actin- and dynamin-dependent manner (see below), we named this new, clathrin-independent internalization route fast endophilin-mediated endocytosis (FEME), as distinct from clathrin-mediated endocytosis (CME).

### Receptor-tyrosine kinase uptake by FEME

As removal of serum (hence of growth factors) from the culture medium or addition of extra serum strongly decreased or increased, respectively, the formation of EPAs (Fig. 2a), we assessed the capacity of various growth factors to activate FEME. Addition of EGF, HGF, FGF, VEGF, PDGF, NGF and IGF-1 but not TGF- $\beta$  or BMP2 induced the formation of EPAs in three different cell lines, including primary fibroblasts from adult donors (Fig. 2a–c, Extended Data Fig. 6a–d and Supplementary Video 4). EGF induced the formation of EPAs in a dose-dependent manner and for prolonged periods of time. Fluorescently

labelled EGF or antibodies specific to the extracellular part of EGFR added to the extracellular medium, but not the  $\beta_1$ -AR, could be detected within EPAs only upon stimulation with EGF (Fig. 2d, e and Extended Data Fig. 6e), establishing that a receptor must be activated by its cognate ligand to become a FEME cargo.



**Figure 2 | Endocytosis of RTKs by the FEME pathway.** **a**, Confocal sections of cells immunostained for endophilin and actin following the indicated treatments. Arrowheads show EPAs. Images are representative of at least ten captures from three independent experiments. **b**, **c**, Quantification of the number of EPAs after 4 min stimulation with various EGF concentrations (**c**) or 10 ng ml<sup>-1</sup> of the indicated growth factors (means  $\pm$  s.e.m. from three independent experiments). **d**, **e**, Co-staining of internalized Alexa555-labelled EGF (**d**, left), anti-EGFR antibody (+EGF; **d**, right) or  $\beta_1$ -AR (+EGF; **e**) and endogenous endophilin. Arrowheads show EPAs (images are representative of ten captures from three independent experiments). **f**, Cell surface staining of endogenous EGFR in resting control, AP2 or endophilin TKD cells (images are representative of ten captures from three independent experiments). Immunoblot shows the amount of EGFR in plasma membrane (PM) fractions of control or endophilin TKD cells. Bar graph shows the quantification of surface EGFR signals at the cell leading edge (LE) in the indicated conditions (means  $\pm$  s.e.m. from three independent experiments; NS, non-significant; \*\*\* $P$  < 0.001, one-way ANOVA and Dunnett's test versus control). **g**, Levels of phosphorylated Erk1/2 in control, AP2 RNAi or TKD cells upon stimulation with 5 ng ml<sup>-1</sup> EGF, determined on immunoblots (means  $\pm$  s.e.m. from three independent experiments; \*\*\* $P$  < 0.001, one-way ANOVA and Dunnett's test versus control  $t$  = 0). **h**, Cell migration into wound sites in samples treated with the indicated RNAi or inhibitors (means  $\pm$  s.e.m. from three independent experiments; \*\*\* $P$  < 0.001, one-way ANOVA and Dunnett's test versus respective control). Scale bars: 20  $\mu$ m (**h**), 10  $\mu$ m (**f**) and 5  $\mu$ m (**a**, **d**, **e**).

Consistent with the binding of endophilin to EGFR and HGFR through CIN85 and Cbl (refs 17,18), depletion of CIN85 or Cbl, or overexpression of a mutant form of CIN85 (which binds Cbl but not endophilin), inhibited EPA formation upon EGF or HGF stimulation but not  $\beta_1$ -AR stimulation (Extended Data Fig. 6f).

EGFR, similarly to  $\beta_1$ -AR, accumulated at the plasma membrane of endophilin TKD cells (Fig. 2f), which resulted in elevated downstream MAPK signalling during the first 10 min of stimulation (Fig. 2g and Extended Data Fig. 7a–g), indicating that endophilin was necessary for maintaining a low basal activation state.

We found that concentrations of EGF over 5 ng ml<sup>-1</sup> were required for full stimulation of EPAs (Fig. 2b), consistent with dynamin-dependent, clathrin-independent endocytic pathway(s) downregulating EGFR signalling at these concentrations<sup>23</sup> (clathrin-mediated endocytosis can function at  $\leq 2$  ng ml<sup>-1</sup> EGF concentrations). However, MAPK signalling was downregulated in endophilin TKD cells upon 30 min of high (100 ng ml<sup>-1</sup>) EGF stimulation (Extended Data Figs 6g–j and 7c), indicating that another endocytic process, probably macropinocytosis, is then mediating EGFR uptake. Notably, FEME carriers and macropinocytosis co-exist but are morphologically distinct—small, <1  $\mu$ m, tubules and vesicles compared with large (2–10  $\mu$ m) ruffles and pinosomes (Extended Data Fig. 6k).

The very prompt budding of endocytic vesicles following receptor stimulation and the enrichment of endophilin at the leading edge suggests that FEME might be involved in directed cell migration. Similarly to EPA formation, migration of endophilin TKD cells was strongly decreased in an EGFR-dependent, but RAF- and MEK-independent, manner (Fig. 2h and Extended Data Fig. 7h), suggesting separate functions for FEME in MAPK signalling and cell migration downstream of EGFR. Finally, consistent with the decrease in NGFR uptake, neurite extension in response to NGF was strongly inhibited in endophilin TKD PC12 cells (Extended Data Fig. 7i–k), consolidating a role for FEME in growth-factor signalling and cell migration.

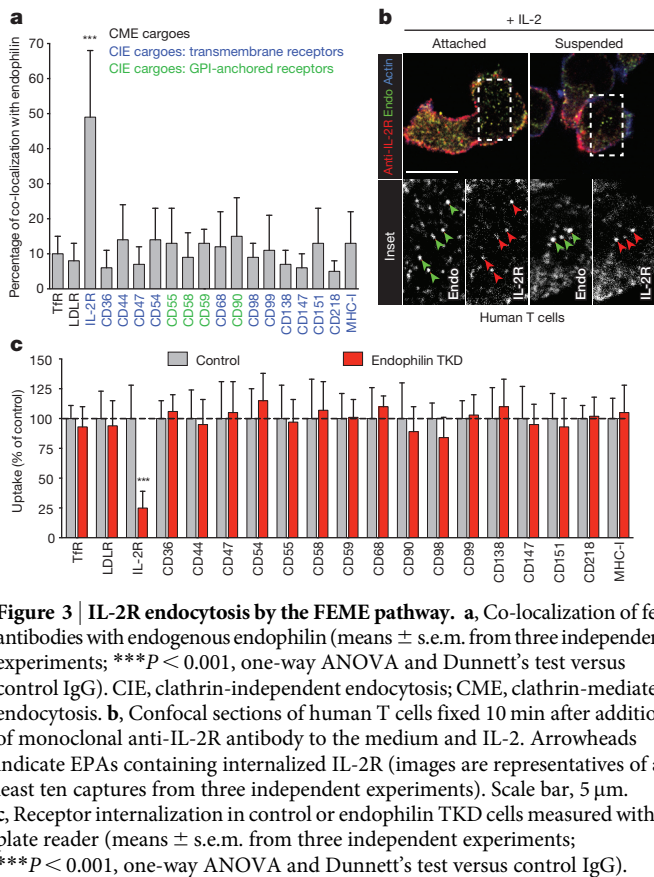
## IL-2 receptor endocytosis by FEME

We next investigated which known clathrin-independent cargoes could use the FEME pathway to enter cells. The  $\beta$ -chain of the interleukin-2 receptor (IL-2R $\beta$ ) and the common cytokine  $\gamma$ -chain (IL-2R $\gamma$ c) are known to use a clathrin-independent endocytic route<sup>24,25</sup>. The IL-2 receptor (composed of the IL-2R $\alpha$ ,  $\beta$  and  $\gamma$ c chains) assembles upon binding to IL-2, triggering a signal transduction cascade leading to lymphocyte proliferation and immune responses. After addition of 200 pM IL-2 together with an antibody specific to the extracellular domain of IL-2R, numerous EPAs containing internalized antibodies were observed in a human T-cell line (Fig. 3a, b). Upon addition of IL-2, endophilin puncta were observed at the edges of cells, followed by numerous internal EPAs (Extended Data Fig. 8a, b). However, we did not detect significant co-localization of endophilin with 16 transmembrane and glycosylphosphatidylinositol (GPI)-anchored receptors of the cluster of designation (CD) superfamily known to be internalized by clathrin-independent mechanisms<sup>26</sup> (Fig. 3a). Consistently, endophilin TKD only decreased the endocytic rate of IL-2R but not that of the other receptors (Fig. 3c). Together, this demonstrated that the FEME pathway mediates the uptake of IL-2-triggered IL-2R but that not all clathrin-independent cargo use this internalization pathway.

## Inhibitors of the FEME pathway

Inhibition of clathrin-mediated endocytosis using five different genetic approaches (RNAi of key components or overexpression of dominant-negative mutants) did not interfere with EPA production upon stimulation (Fig. 4a), further confirming that the FEME pathway acts as a clathrin-independent endocytic route. However, we found that the different chemical means to inhibit clathrin-mediated endocytosis widely used in the literature (hypertonic shock, K<sup>+</sup> depletion, monodansylcadaverine, chlorpromazine and phenylarsine oxide) all strongly inhibited EPA formation, suggesting pleiotropic effects of these inhibitors



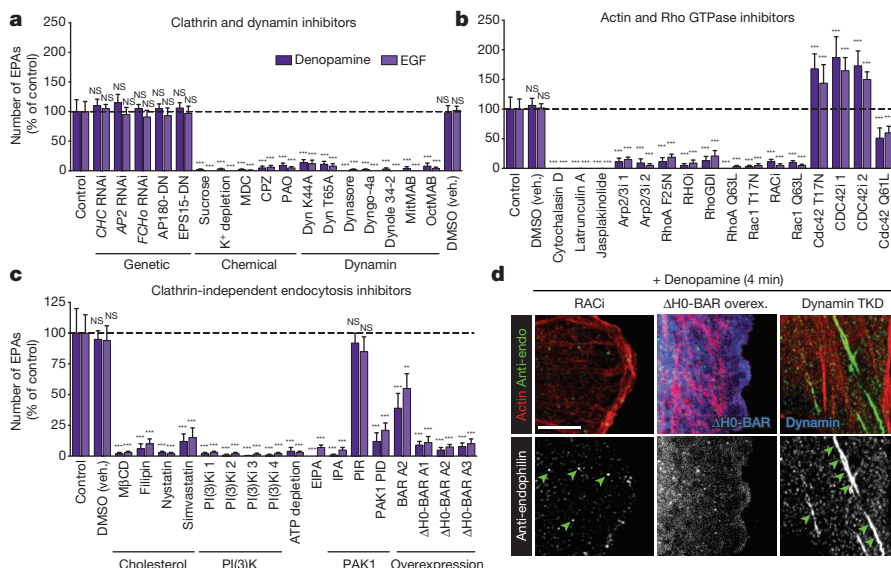


and arguing against their specificity for clathrin-mediated endocytosis. Dynamin is probably the main driver of membrane scission in the FEME pathway, as overexpressing dominant-negative mutants K44A and T65A, depleting its levels by RNAi or inhibiting its function using five different small inhibitors including dynasore and dyngo-4a abrogated EPA budding (Fig. 4a and Extended Data Fig. 8c). Perturbing actin polymerization with cytochalasin D, latrunculin A, jasplakinolide or Arp2/3 inhibitors blocked EPA formation, even at a concentration that did not block clathrin-mediated endocytosis (Fig. 4b and Extended Data Figs 8f, g and 9a). Perturbation of Rho and of Rac blocked the pathway, consistent with their known inhibition of IL-2R uptake<sup>24,27</sup>. Surprisingly, inhibition of Cdc42 increased EPA formation (Fig. 4b,

Extended Data Fig. 8f and Supplementary Video 5). All known inhibitors of clathrin-independent endocytosis/macropinocytosis tested also inhibited the FEME pathway (Fig. 4c). However, overexpression of the BAR domain of endophilin lacking its N-terminal amphipathic helix ( $\Delta$ H0-BAR) was a good inhibitor of the pathway (Fig. 4c, d and Extended Data Fig. 8d, e), suggesting a specific way to inhibit the pathway. It acts by heterodimerizing with endogenous endophilin, and thus decreasing its recruitment to the membrane. Interestingly, many inhibitors acted very promptly (seconds) on endophilin recruitment to the plasma membrane (Extended Data Fig. 8g). Key inhibitors of EPA formation also inhibited FEME cargo uptake but not transferrin uptake and perturbed directed cell migration (Extended Data Fig. 9a, b). Altogether, this established FEME as a clathrin-independent but dynamin-, cholesterol-, actin-, Rac-, Rho-, phosphatidylinositol-3-OH kinase (PI(3)K)- and PAK1-dependent endocytic pathway.

### PtdIns(3,4)P<sub>2</sub> is required for FEME

The prominence of endophilin signals at the leading edge of migrating cells (also called the lamellipodium) and its inhibition by class I PI(3)K inhibitors suggested a role for phosphatidylinositol-3,4,5-triphosphate (PtdIns(3,4,5)P<sub>3</sub>), as it is produced from the phosphorylation of phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) by class I PI(3)K upon receptor activation and is enriched at the leading edge<sup>28,29</sup>. Once generated, PtdIns(3,4,5)P<sub>3</sub> is rapidly hydrolysed back into PtdIns(4,5)P<sub>2</sub> by the 3'-phosphatase PTEN (Fig. 5a and ref. 30). As expected, genetic or chemical inhibition of PTEN significantly increased the recruitment of endophilin at the leading edge in a class I PI(3)K-dependent manner (Fig. 5b, c and Extended Data Fig. 9c). Surprisingly, inhibition of the 5'-phosphatases SHIP1 and SHIP2, that hydrolyse PtdIns(3,4,5)P<sub>3</sub> into PtdIns(3,4)P<sub>2</sub>, strongly reduced the recruitment of endophilin at the leading edge (Fig. 5b, c and Extended Data Fig. 9c). As both the levels of PtdIns(3,4,5)P<sub>3</sub> and PtdIns(3,4)P<sub>2</sub> were elevated in PTEN-inhibited cells (ref. 32 and Extended Data Fig. 9e), but only those of PtdIns(3,4)P<sub>2</sub> were decreased in SHIP1 plus SHIP2-depleted cells, it indicated that PtdIns(3,4)P<sub>2</sub> was required for the recruitment of endophilin. This was confirmed by the unusual recruitment of endophilin onto endosomes/lysosomes in cells depleted for the 4'-phosphatases INPP4A and INPP4B (Extended Data Fig. 9f). INPP4A and INPP4B dephosphorylate PtdIns(3,4)P<sub>2</sub> into PtdIns(3)P at endosomes<sup>33</sup>. We further established that PtdIns(3,4)P<sub>2</sub> produced from PtdIns(4)P by the class II PI(3)K C2 $\alpha$  (which is insensitive to class I PI(3)K inhibitors) was not involved in endophilin recruitment at the leading edge (Fig. 5b, c and Extended Data Fig. 9c), consistent with the recruitment of this kinase to clathrin-coated pits<sup>34</sup>. Thus, we concluded that the formation of



**Figure 4 | Regulation of the FEME endocytic pathway.** a–c, Quantification of the number of EPAs in cells pre-treated with the indicated genetic or chemical clathrin-mediated endocytosis and dynamin inhibitors (CPZ, chlorpromazine; MDC, monodansylcadaverine; PAO, phenylarsine oxide) (a), actin/Rho GTPase inhibitors (b) or clathrin-independent endocytosis inhibitors (EIPA, 5-(N-ethyl-N-isopropyl) amiloride; IPA, IPA-3; MβCD, methyl- $\beta$ -cyclodextrin; PIR, PIR 3.5) (c) before stimulation with denopamine or EGF (means  $\pm$  s.e.m. from three independent experiments; NS, non-significant; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , one-way ANOVA and Dunnett's test versus control). d, Confocal sections of cells depleted for dynamin 1, 2 and 3 (TKD), overexpressing high levels of  $\Delta$ H0-EndoBAR or treated with the EHT1864 (RACi) for 5 min before stimulation with denopamine, fixed and immunostained for endogenous endophilin and actin (images are representatives of at least ten captures). Scale bar, 5  $\mu$ m.





from distinct regions of the cells, on a different timescale to clathrin-mediated endocytosis (Extended Data Fig. 10j). We define the FEME pathway as the rapid formation (seconds in our cells) at the cell surface of tubulo-vesicular (<1 µm) carriers, positive for endophilin, that rapidly travel towards the cell centre promptly after stimulation of cargo receptors by their cognate ligands. Possibly, this will be extended to the growth cones and synapses in neurons where fast clathrin-independent endocytosis was reported<sup>14,15,39</sup>.

Building an endocytic vesicle requires cargo recruitment adaptors, membrane curvature effectors and a membrane scission machinery<sup>40</sup>. Endophilin has all these characteristics in one protein, thus explaining its central role in FEME carrier formation. Its SH3 domain binds to cargo receptors, its BAR domain induces membrane curvature, and by insertion of its multiple amphipathic helices, can support membrane scission<sup>41</sup>, aided by the recruitment of dynamin<sup>13</sup>. We show that, mechanistically, receptors need to be activated to become FEME cargoes, and that PtdIns(3,4)P<sub>2</sub>, produced from the dephosphorylation of PtdIns(3,4,5)P<sub>3</sub> by SHIP phosphatases, mediates the engagement of lamellipodin, which in turn recruits endophilin at the plasma membrane.

The FEME pathway mediates the cellular entry of the well-established clathrin-independent cargoes<sup>24,42,43</sup> cholera and shiga toxins (see accompanying paper<sup>44</sup>) and IL-2R but not of the other cargoes tested. The variety of receptors entering by the FEME pathway suggests a role in many cellular functions and presents ample opportunity for viruses and other pathogens to exploit the pathway.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 15 August; accepted 17 November 2014.**

**Published online 17 December 2014.**

- Ringstad, N., Nemoto, Y. & De Camilli, P. The SH3p4/Sh3p8/SH3p13 protein family: binding partners for synaptotagmin and dynamin via a Grb2-like Src homology 3 domain. *Proc. Natl Acad. Sci. USA* **94**, 8569–8574 (1997).
- Guichet, A. *et al.* Essential role of endophilin A in synaptic vesicle budding at the *Drosophila* neuromuscular junction. *EMBO J.* **21**, 1661–1672 (2002).
- Verstreken, P. *et al.* Endophilin mutations block clathrin-mediated endocytosis but not neurotransmitter release. *Cell* **109**, 101–112 (2002).
- Schuske, K. R. *et al.* Endophilin is required for synaptic vesicle endocytosis by localizing synaptotagmin. *Neuron* **40**, 749–762 (2003).
- Milosevic, I. *et al.* Recruitment of endophilin to clathrin-coated pit necks is required for efficient vesicle uncoating after fission. *Neuron* **72**, 587–601 (2011).
- Ringstad, N. *et al.* Endophilin/SH3p4 is required for the transition from early to late stages in clathrin-mediated synaptic vesicle endocytosis. *Neuron* **24**, 143–154 (1999).
- Gad, H. *et al.* Fission and uncoating of synaptic clathrin-coated vesicles are perturbed by disruption of interactions with the SH3 domain of endophilin. *Neuron* **27**, 301–312 (2000).
- Andersson, F., Low, P. & Brodin, L. Selective perturbation of the BAR domain of endophilin impairs synaptic vesicle endocytosis. *Synapse* **64**, 556–560 (2010).
- Sundborger, A. *et al.* An endophilin-dynamin complex promotes budding of clathrin-coated vesicles during synaptic vesicle recycling. *J. Cell Sci.* **124**, 133–143 (2011).
- Perera, R. M., Zoncu, R., Lucast, L., De Camilli, P. & Toomre, D. Two synaptotagmin 1 isoforms are recruited to clathrin-coated pits at different stages. *Proc. Natl Acad. Sci. USA* **103**, 19332–19337 (2006).
- Ferguson, S. M. *et al.* Coordinated actions of actin and BAR proteins upstream of dynamin at endocytic clathrin-coated pits. *Dev. Cell* **17**, 811–822 (2009).
- Taylor, M. J., Perrais, D. & Merrifield, C. J. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biol.* **9**, e1000604 (2011).
- Meinecke, M. *et al.* Cooperative recruitment of dynamin and BIN/amphiphysin/Rvs (BAR) domain-containing proteins leads to GTP-dependent membrane scission. *J. Biol. Chem.* **288**, 6651–6661 (2013).
- Jockusch, W. J., Praefcke, G. J., McMahon, H. T. & Lagnado, L. Clathrin-dependent and clathrin-independent retrieval of synaptic vesicles in retinal bipolar cells. *Neuron* **46**, 869–878 (2005).
- Llobet, A. *et al.* Endophilin drives the fast mode of vesicle retrieval in a ribbon synapse. *J. Neurosci.* **31**, 8512–8519 (2011).
- Tang, Y. *et al.* Identification of the endophilins (SH3p4/p8/p13) as novel binding partners for the beta1-adrenergic receptor. *Proc. Natl Acad. Sci. USA* **96**, 12559–12564 (1999).
- Petrelli, A. *et al.* The endophilin-CIN85-Cbl complex mediates ligand-dependent downregulation of c-Met. *Nature* **416**, 187–190 (2002).
- Soubeyran, P., Kowanetz, K., Szymkiewicz, I., Langdon, W. Y. & Dikic, I. Cbl-CIN85-endophilin complex mediates ligand-induced downregulation of EGF receptors. *Nature* **416**, 183–187 (2002).
- Cremona, O. *et al.* Essential role of phosphoinositide metabolism in synaptic vesicle recycling. *Cell* **99**, 179–188 (1999).
- Parton, R. G. & del Pozo, M. A. Caveolae as plasma membrane sensors, protectors and organizers. *Nature Rev. Mol. Cell Biol.* **14**, 98–112 (2013).
- Glebov, O. O., Bright, N. A. & Nichols, B. J. Flotillin-1 defines a clathrin-independent endocytic pathway in mammalian cells. *Nature Cell Biol.* **8**, 46–54 (2006).
- Lundmark, R. *et al.* The GTPase-activating protein GRAF1 regulates the CLIC/GEEC endocytic pathway. *Curr. Biol.* **18**, 1802–1808 (2008).
- Sigismund, S. *et al.* Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation. *Dev. Cell* **15**, 209–219 (2008).
- Lamaze, C. *et al.* Interleukin 2 receptors and detergent-resistant membrane domains define a clathrin-independent endocytic pathway. *Mol. Cell* **7**, 661–671 (2001).
- Sauvonnnet, N., Dujancourt, A. & Dautry-Varsat, A. Cortactin and dynamin are required for the clathrin-independent endocytosis of gamma cytokine receptor. *J. Cell Biol.* **168**, 155–163 (2005).
- Maldonado-Báez, L., Williamson, C. & Donaldson, J. G. Clathrin-independent endocytosis: a cargo-centric view. *Exp. Cell Res.* **319**, 2759–2769 (2013).
- Grassart, A., Dujancourt, A., Lazarow, P. B., Dautry-Varsat, A. & Sauvonnnet, N. Clathrin-independent endocytosis used by the IL-2 receptor is regulated by Rac1, Pak1 and Pak2. *EMBO Rep.* **9**, 356–362 (2008).
- Fritsch, R. & Downward, J. SnapShot: Class I PI3K isoform signaling. *Cell* **154**, 940 (2013).
- Servant, G. *et al.* Polarization of chemoattractant receptor signaling during neutrophil chemotaxis. *Science* **287**, 1037–1040 (2000).
- Song, M. S., Salmena, L. & Pandolfi, P. P. The functions and regulation of the PTEN tumour suppressor. *Nature Rev. Mol. Cell Biol.* **13**, 283–296 (2012).
- Xie, J., Erneux, C. & Pirson, I. How does SHIP1/2 balance PtdIns(3,4)P<sub>2</sub> and does it signal independently of its phosphatase activity? *Bioessays* **35**, 733–743 (2013).
- Batty, I. H. *et al.* The control of phosphatidylinositol 3,4-bisphosphate concentrations by activation of the Src homology 2 domain containing inositol polyphosphate 5-phosphatase 2, SHIP2. *Biochem. J.* **407**, 255–266 (2007).
- Gewinner, C. *et al.* Evidence that inositol polyphosphate 4-phosphatase type II is a tumor suppressor that inhibits PI3K signaling. *Cancer Cell* **16**, 115–125 (2009).
- Posor, Y. *et al.* Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. *Nature* **499**, 233–237 (2013).
- Yoon, Y., Zhang, X. & Cho, W. Phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) specifically induces membrane penetration and deformation by Bin/amphiphysin/Rvs (BAR) domains. *J. Biol. Chem.* **287**, 34078–34090 (2012).
- Krause, M. *et al.* Lamellipodin, an Ena/VASP ligand, is implicated in the regulation of lamellipodial dynamics. *Dev. Cell* **7**, 571–583 (2004).
- Vehlow, A. *et al.* Endophilin, Lamellipodin, and Mena cooperate to regulate F-actin-dependent EGF-receptor endocytosis. *EMBO J.* **32**, 2722–2734 (2013).
- Kononenko, N. L. *et al.* Clathrin/AP-2 mediate synaptic vesicle reformation from endosome-like vacuoles but are not essential for membrane retrieval at central synapses. *Neuron* **82**, 981–988 (2014).
- Watanabe, S. *et al.* Ultrafast endocytosis at mouse hippocampal synapses. *Nature* **504**, 242–247 (2013).
- McMahon, H. T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nature Rev. Mol. Cell Biol.* **12**, 517–533 (2011).
- Boucrot, E. *et al.* Membrane fission is promoted by insertion of amphipathic helices and is restricted by crescent BAR domains. *Cell* **149**, 124–136 (2012).
- Römer, W. *et al.* Shiga toxin induces tubular membrane invaginations for its uptake into cells. *Nature* **450**, 670–675 (2007).
- Ewers, H. *et al.* GM1 structure determines SV40-induced membrane invagination and infection. *Nature Cell Biol.* **12**, 11–18 (2010).
- Renard, H.-F. *et al.* Endophilin-A2 functions in membrane scission in clathrin-independent endocytosis. *Nature* <http://dx.doi.org/10.1038/nature14064> (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Y. Peak-Chew for mass spectrometry, M. Edwards and K. McGourty for technical help and P. De Camilli, T. Kirchhausen, G. B. Hammond, P. Robinson, M. Robinson, M. McNiven, B. Nichols, A. Benmerah, M. Krause and Genentech for the gift of reagents, and the members of the McMahon and Boucrot laboratories for comments. The research was funded by the Medical Research Council UK (grant number U105178805) (H.T.M., L.A.-S., G.H., Y.V. and E.B. in part) and a Royal Society grant (research grant number RG120481) (E.B.). A.P.A.F. is supported by the Fundação para a Ciência e Tecnologia, L.A.-S. is a EMBO Long Term fellow and is supported by Marie Curie Actions, and E.B. is a Biotechnology and Biological Sciences Research Council (BBSRC) David Phillips Research Fellow.

**Author Contributions** H.T.M. and E.B. designed the research and supervised the project. H.T.M. and A.P.A.F. performed pull-down experiments. L.A.-S., A.P.A.F. and S.D. did the signalling experiments. L.A.-S. did the super-resolution imaging. A.P.A.F. did the cell migration experiments. S.D. performed the plasma membrane isolations; Y.V. performed the PC12 cell experiments. G.H. prepared and acquired the electron-microscopy data. L.B. and N.S. provided critical reagents. E.B. performed and analysed all the other experiments. E.B. and H.T.M. wrote the manuscript with input from all the other authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.T.M. (hmm@mrclimb.cam.ac.uk) or E.B. (e.boucrot@ucl.ac.uk).



## METHODS

**Cell culture.** BSC1 (ECACC 85011422) and HEK293 cells (ATCC CRL-1573) were cultured in DMEM (Sigma) supplemented with 10% fetal bovine serum (FBS, Gold PAA), 1 mM GlutaMAX-I (Gibco). Normal human epithelial cells hTERT-RPE1 (ATCC CRL-4000) were cultured in DMEM:F12 HAM (1:1 v/v) (Sigma), 0.25% sodium bicarbonate (w/v) (Sigma), 1 mM GlutaMAX-I (Gibco) and 10% FBS (Gold, PAA). Human dermal primary fibroblasts from adult donors (hDFA; ATCC PCS-201-012) were maintained in DMEM:F12 HAM (1:1 v/v) (Sigma), 2% FBS, 1 mM GlutaMAX-I, 0.8  $\mu$ M insulin, 1  $\mu$ g ml<sup>-1</sup> hydrocortisone, 50  $\mu$ g ml<sup>-1</sup> ascorbic acid (Sigma) and 5 ng ml<sup>-1</sup> bFGF (R&D). hDFA were used within 15–20 population doublings *in vitro* to ensure a minimum percentage of senescent cells within the cell population. Human T cells Kit255 (expressing endogenous IL-2R) were cultured in RPMI 1640 (Sigma) supplemented with 10% FBS, 1 mM GlutaMAX-I and 200 pM IL-2. Rat PC12 neuroendocrine cells were cultured in RPMI 1640 supplemented with 7.5% FBS and 7.5% heat inactivated horse serum and 1 mM GlutaMAX-I. BSC1 cells stably expressing  $\sigma$ 2-EGFP were a gift from T. Kirchhausen (Harvard Medical School, USA) and were cultured as parental BSC1 cells with the additional of 200  $\mu$ g ml<sup>-1</sup> G418. Approximately  $2 \times 10^6$ ,  $2 \times 10^5$ ,  $2.5 \times 10^4$  or  $5 \times 10^3$  cells were seeded on 100 mm, 35 mm dishes, 13 mm coverslips or 96-well glass bottom plates, respectively. Cells were regularly tested for mycoplasma contamination.

**Gene cloning.** N-BAR,  $\Delta$ helix0 BAR,  $\Delta$ SH3,  $\Delta$ N-BAR, SH3 domains or full-length endophilin-A1 (SH3GL2, FLJ 92732), -A2 (SH3GL1, IMAGE 3458016) and -A3 iso 1 (SH3GL3, IMAGE 5197246) genes were cloned into pDONR201 (Invitrogen) and transferred into pEGFP, pTagRFP-T<sup>45</sup> (called 'RFP' elsewhere), pTagBFP (called 'BFP' elsewhere) or pGEX-6P2 vectors converted into the Gateway system (pDEST vectors). Mouse  $\beta_1$  adrenergic receptor (ADRB1, IMAGE 905582) full length, third intracellular loop (TIL, amino acids (aa) 246–315) and C terminus (C-term, aa 370–end), human  $\beta_2$  adrenergic receptor (ADRB2, IMAGE 5217922) full length, TIL (aa 220–274) and C-term (aa 330–end), human  $\beta_3$  adrenergic receptor (ADRB3, IMAGE 30915407) TIL (aa 225–294) and C-term (aa 350–end), human  $\alpha_{2a}$  adrenergic receptor (ADRA2A, IMAGE 6198830) full length, TIL (aa 218–374), human  $\alpha_{2b}$  adrenergic receptor (ADRA2B, IMAGE 9020527) full length, TIL (aa 194–372), human dopamine receptor 1 (DRD1, IMAGE 30915297) full length, TIL (aa 219–272) and C-term (aa 338–end), human dopamine receptor 2 (DRD2, IMAGE 3939741) full length, human dopamine receptor 3 (DRD3, IMAGE 40112299) full length, TIL (aa 210–329), mouse dopamine receptor 4 (DRD4, IMAGE 6492837) full length, TIL (aa 309–314), human dopamine receptor 5 (DRD5, IMAGE 3928370) full length, TIL (aa 247–296) and C-term (aa 361–end), human serotonin receptor 6 (HTR6, IMAGE 30915608) TIL (aa 209–265) and C-term (aa 321–end), human serotonin receptor 7 (HTR7, IMAGE 5297325) TIL (aa 258–325) and C-term (aa 389–end), human muscarinic acetylcholine receptor M1 (CHRM1, IMAGE 4931283) full length, TIL (aa 210–366), human muscarinic acetylcholine receptor M2 (CHRM2, IMAGE 40017105) full length, TIL (aa 208–388), human muscarinic acetylcholine receptor M4 (CHRM4, IMAGE 9021535) full length, TIL (aa 217–401), human histamine receptor 1 (HRH1, IMAGE 30340368) TIL (aa 211–418) and human histamine receptor 3 (HRH3, IMAGE 6971899) TIL (aa 218–359), human CIN85 (SH3BP1, IMAGE 3906722) SH3(3) (aa 1–328), PRD (aa 327–428), and PCC (aa 327–end), human Cbl-PRD (aa 477–end), human RhoGDI (ARHGDIA, IMAGE 4867857), rat  $\beta$ -arrestin 1 and 2 full length, human ALIX (PDCD6IP, IMAGE 4340998) PRD domain (aa 717–end) and human PAK1 (PAK1, Addgene 23543) PAK1-inhibitory domain PID (aa 89–149) were cloned into pDONR201 and transferred into pEGFP DEST vectors.

Double-tagged HA- $\beta$ 1AR-EGFP, HA- $\beta$ 2AR-EGFP, HA- $\alpha_{2a}$ AR-EGFP, HA- $\alpha_{2b}$ AR-EGFP, HA-DR1-EGFP, HA-DR2-EGFP, HA-DR3-EGFP, HA-DR4-EGFP, HA-DR5-EGFP, HA-mAChR1-EGFP, HA-mAChR2-EGFP, HA-mAChR4-EGFP plasmids were cloned by inserting a HA tag at the N terminus (thus at the extracellular extremity) into the pDONR201 clones containing full-length receptors described above and then transferred into a pEGFP-C DEST vector. Point mutations  $\beta_1$ -AR R271A, P272/273A, R285A, P283/284A, P272/273/283/284A and R298/299A; DR3 K293E and R318E; DR4 P230/233A, R3236D R237E, R276E; and mAChR4 R230E, K255/256E and P331E were introduced by site-directed mutagenesis in the entry or expression clones containing the TIL of the respective receptors. Alix R757E (ref. 46) was introduced into the PRD clone. Human RhoA F25N, Rac1 T17N, Cdc42 T17N, Rac1 Q61L, RhoA Q63L and Cdc42 Q61L mutants were cloned into pEGFP-expressing vectors. Rat EGFP-LCa (clathrin light chain a) was provided by M. Robinson, rat dynamin 2-EGFP and K44A variant were gifts of M. McNiven, rat EGFP-synaptotagmin 1 (145 and 170 isoforms) and Flag-synaptotagmin 2 (isoform 1) were gifts of P. De Camilli; synaptotagmin 1 D730A (5'-phosphatase mutant<sup>10</sup>) was introduced by quickchange mutagenesis. EGFP-Flotillin-1 was provided by B. Nichols, EGFP-Eps15 A95/295 (called 'Eps15-DN' here) was a gift from A. Benmerah and EGFP-lamellipodin was a gift from M. Krause, the APH domain version of lamellipodin was generated by quickchange mutagenesis. GRAF1-EGFP<sup>22</sup>,

EGFP-AP180 C terminus (residues 530–915, called 'AP180-DN' here)<sup>47</sup>, dynamin1 T65A<sup>48</sup> and FCHO2-EGFP<sup>49</sup> have been described earlier.

**Gene transfection.** For live-cell imaging localization experiments, cells seeded on 35 mm glass bottom dishes (MatTek) were transfected using Lipofectamine 2000 (Invitrogen) or Nanofectin (PAA) and 50–250 ng DNA (the levels of each plasmid were titrated down to levels allowing good detection but limiting side effects of overexpression). For dominant-negative mutants, 1  $\mu$ g of plasmids were used. For endogenous staining on fixed samples, cells seeded on 13 mm coverslips (placed in 24-well plates) were transfected using Lipofectamine 2000 (Invitrogen) or Nanofectin (PAA) and 0.5–50 ng DNA depending on the plasmids and the experiments (low or high overexpression). For receptor tails extracts, cells seeded on 100 mm dishes were transfected using Nanofectin (PAA) and 1–3  $\mu$ g DNA. Cells were incubated 24 h to express the constructs and were either imaged live, fixed (3.7% paraformaldehyde, 20 min) or processed to prepare the cell extracts.

**siRNA suppression of gene expression.** The following siRNA used were: endophilin A1 + 2 + 3 pool 1: Invitrogen Stealth HSS109710, HSS109705 and HSS109713 (targeting human SH3GL2, SH3GL1 and SH3GL3, respectively); endophilin A1 + 2 + 3 pool 2: Invitrogen Stealth HSS109708, HSS109706 and HSS109711 (targeting human SH3GL2, SH3GL1 and SH3GL3, respectively); endophilin A1 + 2 + 3 pool 3: Invitrogen Stealth HSS109709, HSS109707, HSS109712 (targeting human SH3GL2, SH3GL1 and SH3GL3, respectively); endophilin B1 and B2: Dharmacon D-015810-01 and D-015810-02 (targeting human SH3GLB1 and SH3GLB2, respectively); rat EndoA1 + 2 + 3 pool (PC12 cell experiments): 3 oligos against A1 (SH3GL2) RSS3300636, RSS350431 and RSS350432, 3 oligos against A2 (SH3GL1) RSS330960, RSS360961 and RSS372917, 3 oligos against A3 (SH3GL3) RSS330957, RSS330958 and RSS330959; clathrin heavy chain CHC: Invitrogen Stealth HSS174637; AP2: mu2-2 defined in ref. 50 or Invitrogen Stealth HSS101955 (targeting human AP2M1);  $\beta$ -arrestin 1 and 2: Thermo Scientific L-011971 and L-007292 (targeting human ARRB1 and ARRB2, respectively), Cbl: Dharmacon ON-TARGETplus SMARTpool (mix of J-003003-09, J-003003-10, J-003003-11 and J-003003-12 targeting human CBL); CIN85: Dharmacon ON-TARGETplus SMARTpool (mix of J-014748-08, J-014748-07, J-014748-06 and J-014748-05 targeting human SH3KBP1); dynamin 1: Dharmacon ON-TARGETplus SMARTpool (mix of J-003940-05, J-003940-06, J-003940-07 and J-003940-08 targeting human DNM1); dynamin 2: Dharmacon ON-TARGETplus SMARTpool (mix of J-004007-05, J-004007-06, J-004007-07 and J-004007-08 targeting human DNM2); dynamin 3: Dharmacon ON-TARGETplus SMARTpool (mix of J-013931-05, J-013931-06, J-013931-07 and J-013931-08 targeting human DNM3); lamellipodin: Dharmacon ON-TARGETplus SMARTpool (mix of J-031919-08, J-031919-07, J-031919-06 and J-031919-05 targeting human RAPH1); SHIP1: Dharmacon ON-TARGETplus SMARTpool (mix of J-003013-09, J-003013-10, J-003013-11 and J-003013-12 targeting human INPP5D); SHIP2: Dharmacon ON-TARGETplus SMARTpool (mix of J-004152-06, J-004152-07, J-004152-08 and J-004152-09 targeting human INPPL1); PTEN: Dharmacon ON-TARGETplus SMARTpool (mix of J-003023-09, J-003023-10, J-003023-11 and J-003023-12 targeting human PTEN); PI3K-C2 $\alpha$ : Dharmacon ON-TARGETplus SMARTpool (mix of J-006771-08, J-006771-07, J-006771-06 and J-006771-05 targeting human PI3K2CA); INPP4A: Dharmacon ON-TARGETplus SMARTpool (mix of J-011299-06, J-011299-07, J-011299-08 and J-011299-09 targeting human INPP4A); INPP4B: Dharmacon ON-TARGETplus SMARTpool (mix of J-011539-17, J-011539-18, J-011539-19 and J-011539-08 targeting human INPP4B); synaptotagmin 1: Dharmacon ON-TARGETplus SMARTpool (mix of J-019486-07, J-019486-08, J-019486-09 and J-019486-10 targeting human SYNJ1); synaptotagmin 2: Dharmacon ON-TARGETplus SMARTpool (mix of J-019486-07, J-019486-08, J-019486-09 and J-019486-10 targeting human SYNJ2). Control siRNA used were Dharmacon ON-TARGETplus GAPDH control (D-001830-01-05), Invitrogen Stealth control (scrambled) oligo 138782 or a mixture of Invitrogen Stealth 'high GC' 45-2000 and 'low GC' 45-2002 oligos.

Cells seeded on 35 mm dishes were transfected twice (on day 1 and 2) with Oligofectamine (Invitrogen) complexed with 80 pmol of each indicated siRNA and analysed 3–4 days after the first transfection. Rescue experiments were performed with plasmid DNA transfection of the indicated constructs at the middle of day 3 (when endogenous levels of the targeted proteins were already decreased) and the cells were analysed 16 h later (on day 4). Transfection of siRNA into human T cells Kit255 was performed by electroporation using a Nucleofector 2b (Lonza) and Ingenio (Mirus Bio) cuvettes and electroporation solutions. RNAi knockdown efficiency was verified by western blotting and immunofluorescence counter staining, when antibodies were available. The use of validated pools of siRNA targeting the same genes increased the knock-down efficiency and specificity.

**Protein purification and pull-down experiments.** GST-tagged constructs were expressed in BL21 *Escherichia coli* (Stratagene). Cells were lysed using an Emulsiflex C3, spun at 140,000g for 40 min at 4 °C in a Beckman Ti45 rotor, and the supernatant was bound to glutathione beads for 30 min. The beads were washed extensively with 150 mM NaCl, 20 mM HEPES pH 7.4, 2 mM DTT, 2 mM EDTA, including

2 washes with 500 mM NaCl. GST-proteins were eluted from the GST-sepharose beads with 10 mM glutathione, further purified by Superdex 200 gel filtration, and rebound to a minimal volume of fresh GST-sepharose beads (to achieve saturation) for use in pull downs. HEK293 cells expressing various EGFP-tagged constructs were quickly washed with cold PBS, lysed in ice-cold lysis buffer (150 mM NaCl, 20 mM HEPES, 5 mM DTT, 0.1% Triton X-100 and a protease and phosphatase inhibitor cocktail (Thermo Scientific)), briefly sonicated and spun at 14,000g for 10 min. Bead-bound proteins were then exposed to cell lysates for 1 h, pelleted in a cooled bench-top centrifuge, and washed three times in lysis buffer. The final pellet was boiled in sample buffer and run on SDS-PAGE ('input' lanes correspond to 10% of cell extract). The proteins were transferred onto PVDF membrane and immunoblotted using anti-EGFP antibodies.

**Antibodies.** The following antibodies were used for immunostaining, flow cytometry or immunoblotting: pan anti-endophilin Ra74 (in-house affinity purified rabbit polyclonal raised against the SH3 domain of rat endophilin A1<sup>15</sup>), anti-endophilin A2 clones H-60 and A-11 (rabbit polyclonal sc-25495 and mouse monoclonal sc-365704, respectively, Santa Cruz Biotechnology), anti-EGFR antibody clone D38B1 used for total staining (rabbit polyclonal, Cell Signaling Technologies 4267), anti- $\alpha$ -adaptin clone 8 (mouse monoclonal, BD Bioscience 610501) and clone AP6 (mouse monoclonal, Thermo Scientific MA1-064), anti-clathrin heavy chain X22 (Thermo Scientific MA1-065), anti-vinculin clone hVIN-1 (mouse monoclonal, Sigma V9264), anti-Arp3 (rabbit polyclonal, Millipore 07-277), anti-cortactin p80/85 clone 4F11 (mouse monoclonal, Millipore 05-180), anti-dynamin clone 41 (mouse monoclonal, BD Transduction Laboratories 610245), anti-caveolin-1 (mouse monoclonal, BD Transduction Laboratories 610059), anti-lamellipodin (mouse monoclonal clone H-5, Santa Cruz sc-390050 and rabbit polyclonal Atlas Antibodies HPA020027), anti-SHIP1 clone D1163 (rabbit polyclonal, Cell Signaling Technologies 2728), anti-SHIP2 clone C76A7 (rabbit polyclonal, Cell Signaling Technologies 2839), anti-synaptojanin (rabbit polyclonal, Abcam 84309), anti-LAMP-1 clone H4A3 (Developmental Studies Hybridoma Bank), anti-calnexin H60 (rabbit polyclonal, Santa Cruz Biotechnology sc-11397), anti-integrin  $\beta$ 1 (rabbit polyclonal, Cell Signaling Technologies 4706), anti- $\beta$ -AR (rabbit polyclonal, Cell Signaling Technologies 12271), anti- $\beta$ -arrestin1/2 (goat polyclonal, Abcam 31294), anti-phosphorylated p44/42 MAPK T202/Y204 (phosphorylated Erk1/2) clone D13.14.4E (rabbit polyclonal, Cell Signaling Technologies 4370), anti-p44/42 MAPK (total Erk1/2) (rabbit polyclonal, Cell Signaling Technologies 9102), anti-phosphorylated Elk1 (Ser 383) (rabbit polyclonal, Cell Signaling Technologies 9181), anti-phosphorylated cJun (Ser 73) clone D47G9 (rabbit polyclonal, Cell Signaling Technologies 3270), anti-phosphorylated CREB (Ser 133) clone 87G3 (rabbit polyclonal, Cell Signaling Technologies 9198), anti-PtdIns(3,4)P<sub>2</sub> (mouse monoclonal, Echelon Bioscience Z-P034b), anti-PtdIns(4,5)P<sub>2</sub> (mouse monoclonal, Echelon Bioscience Z-A045), anti-PtdIns(3,4,5)P<sub>3</sub> (mouse monoclonal, Echelon Bioscience Z-P345b), anti- $\beta$ -tubulin clone 9F3 (mouse monoclonal, Cell Signaling Technologies 2128), anti-EGFP ab290 (rabbit polyclonal, AbCam 290), anti-Myc tag clone 9E10 (mouse monoclonal, AbGent AM1007a) and clone 9B11 (mouse monoclonal, Cell Signaling Technologies 2276), anti-Flag tag clone M2 (mouse monoclonal, Sigma F1804 and rabbit polyclonal Cell Signaling Technologies 2368), anti-GAPDH clone 0411 (mouse monoclonal, Santa Cruz Biotechnology sc-47724) and clone 14C10 (rabbit polyclonal, Cell Signaling Technologies 2118).

The following antibodies were used for receptor uptake assays ('antibody feeding assays'): anti-HA clone 16B12 used for 'HA-GPCR-EGFP' chimera uptake assays (mouse monoclonal Covance), anti-EGFR antibody clone 13A9 (mouse monoclonal raised against the ectodomain of EGFR that does not compete with EGF binding, a gift from Genentech), anti-IL-2R clone 561 (mouse monoclonal, recognizing the extracellular portion of IL-2R but does not compete with IL-2 (ref. 51)), anti-NGFR antibody clone 165131 (mouse monoclonal raised against the ectodomain of TrkA, R&D systems MAB2148), anti-TfR clone CBL47 (rabbit polyclonal, Millipore CBL47), anti-LDLR (mouse monoclonal, R&D systems MAB2148), anti-CD36 clone 5-271 (mouse monoclonal, BioLegend 336202), anti-CD44 clone BJ18 (mouse monoclonal, BioLegend 338802), anti-CD47 clone CC2C6 (mouse monoclonal, BioLegend 323102), anti-CD54 (also called ICAM-1) clone HA58 (mouse monoclonal, BioLegend 353101), anti-CD55 clone JS11 (mouse monoclonal, BioLegend 311302), anti-CD58 clone LFA-3 (mouse monoclonal, BioLegend 330902), anti-CD59 clone H19 (mouse monoclonal, BioLegend 304702), anti-CD68 clone Y82A (mouse monoclonal, BioLegend 333801), anti-CD90 (Thy1) clone IBL-6/23 (rat monoclonal, AbCam ab3105), anti-CD98 clone MEM-108 (mouse monoclonal, BioLegend 315602), anti-CD99 clone TU12 (mouse monoclonal, BD Pharmingen 555687), anti-CD138 clone MI15 (mouse monoclonal, BioLegend 356501), anti-CD147 clone HIM6 (mouse monoclonal, BioLegend 306202), anti-CD151 (mouse monoclonal, ABD Serotec MCA1856T), anti-CD218a clone H44 (mouse monoclonal, BioLegend 313804) and anti-MHC-I HLA-A,B,C clone W6/32 (mouse monoclonal, BioLegend 311402).

Primary antibodies were used at 0.2–5  $\mu$ g ml<sup>-1</sup> for immunoblotting, 0.1–10  $\mu$ g ml<sup>-1</sup> for immunofluorescence and 10  $\mu$ g ml<sup>-1</sup> for antibody feeding assays. Before immunostaining, fixed cells were permeabilized with 0.05% saponin, apart from staining of activated transcription factors (CREB, Elk1 and Jun) which was done after permeabilization with 0.1% Triton X-100 and for phosphoinositide staining which was done as described below.

The following secondary antibodies were used: Alexa Fluor 488 and Alexa Fluor 555 goat anti-mouse IgG, Alexa Fluor 488 and Alexa Fluor 555 goat anti-rabbit IgG, Alexa Fluor 488 goat anti-rat IgG, Alexa Fluor 488 donkey anti-goat IgG and Alexa Fluor 555 donkey anti-rabbit (all from Life Technologies) and goat anti-mouse IgG-HRP conjugate and goat anti-rabbit IgG-HRP conjugate (both from Bio-Rad).

**Chemicals, growth factors and small inhibitors.** Pre-treatments with small inhibitors before stimulation were performed for 5 min at 37 °C, unless specified otherwise, and the inhibitors were kept during the stimulation or assay periods in all cases. Cytochalasin D (actin depolymerizer, Tocris 1233, used at 100 nM), latrunculin A (actin depolymerizer, Tocris 3973, used at 200 nM), jasplakinolide (actin stabilizer, Tocris 2792, used at 100 nM), rhosin (RhoA inhibitor called 'RHOi' in this study, Calbiochem 555460, used at 10  $\mu$ M), EHT1864 (RAC inhibitor, called 'RACi 1' in this study, Tocris 3872, used at 2  $\mu$ M), EHOP-016 (RAC inhibitor, called 'RACi 2' in this study, Sigma SML0526, used at 2  $\mu$ M), secramine<sup>52</sup> (Cdc42 inhibitor, called 'CDC42i 1' in this study, a gift from T. Kirchhausen (Harvard Medical School) and the Hammond lab (Univ. of Louisville) was synthesized by B. Xu and G. B. Hammond, and was used at 10  $\mu$ M), ML141 (Cdc42 allosteric inhibitor, called 'CDC42i 2' in this study, Tocris 4266, used at 10  $\mu$ M), CK548 (Arp2/3 inhibitor, called 'Arp2/3i 1' in this study, Sigma C7499, used at 50  $\mu$ M), CK636 (Arp2/3 inhibitor, called 'Arp2/3i 2' in this study, Sigma C7374, used at 50  $\mu$ M), methyl- $\beta$ -cyclodextrin (extracts cholesterol, called 'M $\beta$ CD' in this study, Sigma 332615, used at 50  $\mu$ M for three incubations of 10 min each), filipin (sequesters cholesterol, Sigma F4767, used at 1  $\mu$ g ml<sup>-1</sup>), nystatin (sequesters cholesterol, Sigma N3503, used at 10  $\mu$ g ml<sup>-1</sup>), simvastatin (sequesters cholesterol, Sigma S6196, used at 10  $\mu$ M), dynasore (dynamin inhibitor<sup>53</sup>, a gift of T. Kirchhausen (Harvard Medical School) and used at 80  $\mu$ M), Dyngo-4a (dynamin inhibitor<sup>54</sup>, a gift of P. Robinson (Univ. of Sydney) and used at 20  $\mu$ M), dynole 34-2 (dynamin inhibitor<sup>55</sup>, Tocris 4222, used at 20  $\mu$ M), MitMAB (dynamin inhibitor<sup>56</sup>, Calbiochem 324411, used at 20  $\mu$ M), OctMAB (dynamin inhibitor<sup>56</sup>, Tocris 4225, used at 20  $\mu$ M), monodansylcadaverine (clathrin-mediated endocytosis inhibitor<sup>57</sup>, called 'MDC' in this study, Sigma D4008, used at 100  $\mu$ M), chlorpromazine (clathrin-mediated endocytosis inhibitor<sup>58</sup>, called 'CPZ' in this study, Sigma C8138, used at 10  $\mu$ g ml<sup>-1</sup>), phenylarsine oxide PAO (clathrin-mediated endocytosis inhibitor<sup>59</sup>, Sigma P3075, used at 10  $\mu$ M), GDC-0941 (broad and specific PI(3)K inhibitor, called 'PI(3)Ki 1' or 'PI(3)Ki' in this study, Symansis SYGDC0941, used at 50 nM), GSK2126458 (broad and specific PI(3)K inhibitor, called 'PI(3)Ki 2' in this study, BioVision 1961, used at 25 nM), LY294002 (broad PI(3)K inhibitor, called 'PI(3)Ki 3' in this study, Tocris 1130, used at 50  $\mu$ M), wortmannin (broad PI(3)K inhibitor, called 'PI(3)Ki 4' in this study, Cell Signaling Technologies 9951, used at 500 nM), A66 S (PI(3)K p110 $\alpha$  specific inhibitor, Selleckchem S2636, used at 50 nM), TGX-221 (PI(3)K p110 $\beta$  specific inhibitor, Cayman 10007349, used at 25 nM), CAL-101 (PI(3)K p110 $\delta$  specific inhibitor, Selleckchem S2226, used at 50 nM), AS-252424 (PI(3)K p110 $\gamma$  specific inhibitor, Cayman 10009052, used at 100 nM), 5-(N-ethyl-N-isopropyl) amiloride (macropinocytosis inhibitor, called 'EIPA' in this study, Sigma A3085, used at 25  $\mu$ M), PF-3758309 (broad PAK inhibitor, called 'PAKi' in this study, Calbiochem 500613, used at 1  $\mu$ M), IPA-3 (PAK1 inhibitor, called 'IPA' in this study, Sigma I2285, used at 20  $\mu$ M) and PIR 3.5 (IPA-3 inactive analogue, called 'PIR' in this study, Tocris 4212, used at 20  $\mu$ M), AS 19499490 (SHIP2 inhibitor, called 'SHIP2i' in this study, Tocris 3718, used at 1  $\mu$ M), bpV(Hop) (protein tyrosine phosphatase inhibitor with selectivity for PTEN (IC<sub>50</sub> = 14 nM), called 'PTENi 1' in this study, Cayman 14433, used at 200 nM), bpV(Phen) (protein tyrosine phosphatase inhibitor with selectivity for PTEN (IC<sub>50</sub> = 38 nM), called 'PTENi 2' in this study, Santa Cruz Biotechnology sc-221378, used at 200 nM), SF1670 (PTEN inhibitor, called 'PTENi 3' in this study, Cayman 15368, used at 10  $\mu$ M), PD153035 (EGFR inhibitor, called 'EGFRi' in this study, Tocris 1037, used at 5  $\mu$ M), SB590885 (RAF inhibitor, called 'RAFi' in this study, Tocris 2650, used at 5  $\mu$ M), PD032591 (MEK inhibitor, called 'MEKi 1' in this study, Tocris 4192, used at 0.5 mM), U0126 (MEK inhibitor, called 'MEKi 2' in this study, Cell Signaling Technologies 99035, used at 10 mM). Sucrose hypertonic treatment (inhibiting clathrin-mediated endocytosis<sup>60</sup>) was performed by incubating cells at 37 °C with a 0.45 M sucrose solution; potassium (K<sup>+</sup>) depletion (inhibiting clathrin-mediated endocytosis<sup>61</sup>), was performed by washing the cells (all at 37 °C) with PBS twice, followed by 5 min with hypotonic PBS (1:3 v/v with water), and five quick washes followed by 10 min incubation in K<sup>+</sup>-depletion solution (150 mM NaCl, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 2 mM HEPES pH 7.4, 0.5% DMSO, 0.5% BSA). Cellular ATP depletion was performed as in ref. 42 by washing cells twice followed by incubation for 30 min at 37 °C in 10 mM



2-deoxy-D-glucose, 10 mM Na<sub>3</sub>, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub> in PBS. Control sample was prepared in 5 mM glucose, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub> in PBS.

The following agonists, antagonists and growth factors were used: denopamine ( $\beta_1$ -AR selective agonist, Sigma D7815), isoproterenol (non-selective  $\beta$ -AR agonist, Sigma I5627), noradrenaline ( $\beta_1$ -AR agonist, Sigma, A7257), dobutamine ( $\beta_1$ -AR partial agonist, Sigma, D0676), terbutaline ( $\beta_2$ -AR selective agonist, Sigma, T2528), atenolol ( $\beta_1$ -AR selective antagonist, Sigma A7655), betaxolol ( $\beta_1$ -AR selective antagonist, Sigma B5683), butoxamine ( $\beta_2$ -AR selective antagonist, Sigma B1385), clonidine ( $\alpha_2$ -AR selective agonist, Sigma C7897), dexmedetomidine ( $\alpha_2$ -AR selective agonist, Tocris 2749), dopamine (dopamine receptor agonist, Sigma H8502), quinpirole (dopamine receptor 2 agonist, Sigma Q102), (+)-PD128907 (dopamine receptor 3 selective agonist, Tocris 1243), PD168077 (dopamine receptor 4 selective agonist, Tocris 1065), acetylcholine (muscarinic acetylcholine receptor agonist, Sigma A6625), VU152100 (muscarinic acetylcholine receptor 4 selective agonist, Tocris 3383), epidermal growth factor (EGF, Sigma E9644), Alexa Fluor 555-EGF (Life Technologies E35350), hepatocyte growth factor (HGF, Calbiochem 375228), fibroblast growth factor (FGF, Calbiochem 341591), vascular endothelial growth factor (VEGF, Calbiochem 676473), platelet-derived growth factor (PDGF, Calbiochem 521215), insulin-like growth factor 1 (IGF-1, Calbiochem 407251), transforming growth factor beta 1 (TGF- $\beta_1$ , Sigma T5050), bone morphogenetic protein 2 (BMP2, Cell Signaling Technologies 4697), nerve growth factor (NGF 2.5S, Invitrogen 13257-019), holo-transferrin (Sigma T0665), Alexa Fluor 488-Transferrin (Life Technologies T13342), Alexa Fluor 546-Transferrin (Life Technologies T23364), Alexa Fluor 633-Transferrin (Life Technologies T23362) and DiL-low density lipoprotein (DiL-LDL, Life Technologies L3482). Growth factors were labelled with Alexa-555 using a protein labelling kit (Life Technologies A37571), according to the instructions from the manufacturer.

**Cell stimulation and cargo uptake.** Cells were never serum-starved before stimulation and the ligand and antibodies were never pre-incubated with the cells at 4 °C. 'Resting' conditions correspond to cells being cultured in 10% serum media and directly fixed. '+ Serum' stimulation corresponds to the addition of 10% serum to regular culture media (final 20% serum concentration). Specific stimulations were performed as follow. HA-EGFP double tagged GPCRs: cells expressing HA-EGFP double-tagged GPCRs were stimulated at 37 °C, 5% CO<sub>2</sub> for 10 min in the presence of 10  $\mu\text{g ml}^{-1}$  of anti-HA clone 16B12 monoclonal antibody and 10  $\mu\text{M}$  of the ligands denopamine (HA- $\beta_1$ AR-EGFP), isoproterenol (HA- $\beta_2$ AR-EGFP), clonidine (HA- $\alpha_{2a}$ AR-EGFP and HA- $\alpha_{2b}$ AR-EGFP), dopamine (HA-DR1-EGFP and HA-DR5-EGFP), quinpirole (HA-DR2-EGFP), (+)-PD128907 (HA-DR3-EGFP), PD168077 (HA-DR4-EGFP), acetylcholine (HA-muR1-EGFP and HA-muR2-EGFP) or VU152100 (HA-muR4-EGFP). For microscopy experiments, cells expressing HA- $\beta_1$ -AR-EGFP were incubated at 37 °C, 5% CO<sub>2</sub> for 4 min with 10  $\mu\text{g ml}^{-1}$  anti-HA clone 16B12 monoclonal antibody and 10  $\mu\text{M}$  denopamine. Adrenergic agonists and antagonists: cells were incubated at 37 °C for 4 min with 10  $\mu\text{M}$  or a titrating range ( $10^{-8}$  to  $10^{-5}$  M) of concentration of denopamine, isoproterenol, noradrenaline, dobutamine or terbutaline. In some experiments, cells were incubated with 10  $\mu\text{M}$  atenolol, betaxolol, butoxamine 37 °C for 5 min before stimulation with isoproterenol. Growth factors: cells were incubated at 37 °C for 4 min with pre-warmed serum-free media containing 10 ng ml<sup>-1</sup> of EGF, HGF, FGF, VEGF, PDGF, IGF-1 or NGF or 50 ng ml<sup>-1</sup> BMP2 or 5 ng ml<sup>-1</sup> TGF- $\beta_1$  or 50 ng ml<sup>-1</sup> Alexa Fluor 555-conjugated EGF, HGF, FGF, VEGF or IGF-1. In some experiments, EGF was titrated from 2 to 500 ng ml<sup>-1</sup>. Anti-EGFR antibody feeding assays were performed by directly incubating cells with pre-warmed serum-free media containing 0.5  $\mu\text{g ml}^{-1}$  of anti-EGFR clone 13A9 monoclonal antibody and 50 ng ml<sup>-1</sup> of the indicated growth factors or 10  $\mu\text{M}$  denopamine for 4 min at 37 °C. Interleukin-2: human T cells Kit255 were incubated at 37 °C with pre-warmed serum-free media containing either 200 pM IL-2 or 10  $\mu\text{g ml}^{-1}$  anti-IL-2R clone 561 antibody and 200 pM IL-2 for the indicated time. Clathrin-independent cargoes: cells were incubated at 37 °C for 15 min with pre-warmed serum-free media containing 10  $\mu\text{g ml}^{-1}$  of the indicated anti-CD proteins or MHC-I antibodies detecting the extracellular portion of the receptors. Clathrin-dependent cargoes: cells were incubated at 37 °C for 7 min with pre-warmed serum-free media containing 20  $\mu\text{g ml}^{-1}$  Alexa Fluor-488-labelled human transferrin or 50  $\mu\text{g ml}^{-1}$  DiL-LDL or for 15 min with media containing 10  $\mu\text{g ml}^{-1}$  anti-TfR and anti-LDLR antibodies.

**Sample processing and analysis.** *Samples for microscopy.* After the incubation period at 37 °C, cells stimulated as described above were quickly washed once with 37 °C pre-warmed PBS to removed unbound ligands and antibodies and fixed with pre-warmed PFA 3.7% for 20 min at 37 °C. Fixed cells were then washed three washes three times with PBS and one time with PBS supplemented with 50 mM NH<sub>4</sub>Cl to quench free PFA. Cells were then permeabilized (0.05% saponin), immunostained and imaged as described below.

*Samples to be analysed by plate reader.* After the incubation period at 37 °C, cells stimulated as described above were placed on ice to stop membrane traffic, washed with ice-cold PBS to removed unbound antibodies and incubated with

Alexa-488-conjugated secondary antibodies on ice to label non internalized antibodies bound to receptors remaining on the cell surface. Cells were then washed twice with PBS, and fixed on ice with 3.7% PFA for 30 min. Cells were washed twice with PBS and once with 50 mM NH<sub>4</sub>Cl in PBS and incubated with DAPI to label their DNA (used to normalized the number of cells in between samples and experiments). Signals from remaining (non-internalized) receptors and DAPI were measured using a FLUOstar Optima (BMG) plate reader.

*Samples to be analysed by flow cytometry.* After the incubation period at 37 °C, cells stimulated as described above were then placed on ice to arrest membrane traffic, washed with ice-cold PBS to removed unbound ligands, detached by 1 min incubation with 0.25% Trypsin-EDTA at 37 °C, spun, acid-washed (ice-cold stripping buffer (150 mM NaCl, 5 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, glycine) pH 2.5, three washes to remove surface-bound ligands), washed with PBS twice (to restore pH 7.4), fixed (PFA 3.7% for 20 min), washed and re-suspended in PBS and analysed using LSR II flow cytometer (Beckson-Dickinson).

**Cell surface labelling.** Cell surface staining of HA-GPCR-EGFP and EGFR was performed by incubating cells at 4 °C for 1 h with 10  $\mu\text{g ml}^{-1}$  of anti-HA clone 16B12 monoclonal antibody or 10  $\mu\text{g ml}^{-1}$  anti-EGFR 13A9 monoclonal antibody respectively, diluted in PBS containing 5% horse serum, followed by a quick wash with cold PBS and fixation (PFA 3.7% for 30 min at 4 °C followed by three washes with PBS and one wash with PBS supplemented with 50 mM NH<sub>4</sub>Cl). Cells were then incubated with secondary antibodies diluted in 5% horse serum, and the coverslips were mounted with DABCO anti-fade agent on glass slides and imaged using the laser scanning confocal microscope described below. Signals were imaged by confocal microscopy or surface levels were measured using a FLUOstar Optima (BMG) plate reader.

**Immunostaining, laser scanning confocal fluorescent microscopy and analysis.** For ligand-induced endophilin-positive assemblies (EPAs), budding cells pre-incubated or not with inhibitors for 5 min were incubated at 37 °C for 4 min with pre-warmed imaging buffer containing the indicated ligands (note that no serum starvation or pre-incubations at 4 °C were performed), quickly rinsed with pre-warmed PBS and fixed (PFA 3.7% for 20 min at 37 °C followed with three washes with PBS and one wash with PBS supplemented with 50 mM NH<sub>4</sub>Cl). Cells were then permeabilized (0.05% saponin) and immunostained (primary and secondary antibodies were diluted in 5% horse serum). In some experiments, DRAQ5 (DNA-binding dye, Cell Signaling Technologies 4084), phalloidin-tetramethylrhodamine (Santa Cruz Biotechnology sc-362065) or phalloidin-AlexaFluor 647 (Cell Signaling Technologies 8940) were added with secondary antibodies. The coverslips were mounted with DABCO anti-fade agent on glass slides and imaged using a laser scanning confocal microscope (TCS Sp5 AOBs; Leica) equipped with a 63 $\times$  objective. For Alexa488, the illumination was at 488 nm and emission collected between 498 and 548 nm; for Alexa555 the laser illumination was at 543 nm and emission collected between 555 and 620 nm; for Alexa647 and DRAQ5, the laser illumination was at 633 nm and emission collected between 660 and 746 nm. For each quantification, the number of EPAs within 400  $\mu\text{m}^2$  masks were counted on confocal planes placed at the middle of three representative cells (from independent experiments). The percentages of co-localization of markers on EPAs or endophilin spots located at the leading edge of cells were determined using Volocity 6.0. The levels of endophilin (endogenous and EGFP-tagged constructs) and lamellipodin at the leading edges of cells were measured, for each condition, on masks covering a total at least 100  $\mu\text{m}$  long ( $\sim 2 \mu\text{m}$  wide) of leading edge of three cells from independent experiments using Volocity 6.0.

**Immunostaining and plate reader measurements.** Resting or stimulated cells grown on glass bottom 96-well plates were placed on ice to stop membrane traffic, washed with PBS to removed unbound ligands and fixed with 3.7% PFA for 30 min. Cells were washed twice with PBS and once with 50 mM NH<sub>4</sub>Cl in PBS and permeabilized and immunostained with primary and secondary antibodies together with DAPI to label DNA (used to normalize the number of cells in between samples and experiments). Signals were measured using a FLUOstar Optima (BMG) plate reader.

**Phosphoinositides staining and measurements.** Cells grown in glass-bottom 96-well plates were fixed for 1 h at 37 °C with 3.7% (w/v) PFA, 0.1% glutaraldehyde and 0.15 mg ml<sup>-1</sup> saponin in 5 mM KCl, 137 mM NaCl, 4 mM NaHCO<sub>3</sub>, 0.4 mM KH<sub>2</sub>PO<sub>4</sub>, 1.1 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM MgCl<sub>2</sub>, 5 mM PIPES, pH 7.2, 2 mM EGTA and 5.5 mM glucose, as in ref. 62. Samples were then incubated for 90 min with primary antibodies diluted in 1% BSA, washed, stained with secondary antibodies together with DAPI to label DNA (used to normalize the number of cells in between samples and experiments). Signals were measured using a FLUOstar Optima (BMG) plate reader. The specificity of the antibodies was verified on cells overexpressing the PtdIns(4,5)P<sub>2</sub>, PtdIns(3,4)P<sub>2</sub> and PtdIns(3,4,5)P<sub>3</sub>-specific PH domains of PLC $\gamma$ , TAPP1 and BTK, respectively.

**Super-resolution microscopy.** Cells were grown on 13 mm #1.5 Hecht Assistant coverslips and processed for immunofluorescence as described above with a few



modifications: for structured illumination microscopy (SIM) we used twice the amount of primary antibodies while for stimulated emission depletion microscopy (STED) both primary and secondary antibodies were used at five times its recommended concentration. Samples were mounted using ProLong gold (Life Technologies). SIM images were acquired in Zeiss Elyra S.1 and processed using the Zeiss acquisition software ZEN. STED images were acquired in a Leica TCS SP8 X gated STED microscope equipped with a 592 depletion laser.

**Immuno-electron microscopy.** Control and stimulated (10  $\mu$ M denopamine, 4 min at 37 °C) BSC1 cells grown on MatTek glass-bottomed Petri dishes were fixed in either 4% paraformaldehyde alone or with the addition of 0.05% glutaraldehyde in 0.1 M phosphate buffer, pH 7.4 overnight at 4 °C. After several buffer washes reactive aldehyde groups were inactivated using 0.1% sodium borohydride in phosphate buffer for 30 min, followed by permeabilization with 0.03% saponin in 20 mM phosphate buffer, 150 mM sodium chloride. Cells were then incubated in normal goat serum (Aurion) for 40 min before incubation with 0.5  $\mu$ g ml<sup>-1</sup> antibody against endophilin for 4.5 h at room temperature. After thorough washing, cells were incubated in a 1:200 dilution of goat anti rabbit ultrasmall gold (Aurion) overnight at 4 °C. Cells were then fixed in 2% glutaraldehyde in phosphate buffer, washed with distilled water followed by silver enhancement of gold using R-Gent SE-EM (Aurion) reagents and post fixed with 0.5% osmium tetroxide in 0.1 M phosphate buffer on ice for 15 min. Cells were dehydrated in an ascending ethanol series and embedded in CY212 resin. Ultrathin sections were stained with saturated aqueous Uranyl Acetate and Reynolds Lead Citrate and examined using an FEI Technai Spirit EM operated at 80 kV.

**Live-cell confocal fluorescent microscopy and analysis.** Just before live-cell imaging, the medium of cells grown on MatTek dishes was changed to  $\alpha$ -MEM without phenol red, supplemented with 20 mM HEPES, pH 7.4 and 5% FBS and placed into a temperature controlled chamber on the microscope stage with 95% air: 5% CO<sub>2</sub> and 100% humidity. Live-cell imaging data were acquired using a fully motorized inverted microscope (Eclipse TE-2000, Nikon) equipped with a CSU-X1 spinning disk confocal head (UltraVIEW VoX, Perkin-Elmer, England) using a 60 $\times$  lens (Plan Apochromat VC, 1.4 NA, Nikon) under control of Velocity 6.0 (Improvision, England). 14-bit digital images were obtained with a cooled EMCCD camera (9100-02, Hamamatsu, Japan). Four 50 mW solid-state lasers (405, 488, 561 and 647 nm; Crystal Laser and Melles Griots) coupled to individual acoustic-optical tunable filter (AOTF) were used as light source to excite EGFP, Alexa488, TagRFP-T, Alexa555 and 647, as appropriate. Rapid two-colour time-lapses were acquired at 500 ms to 2 s intervals, using a dual (525/50; 640/120, Chroma) emission filter respectively. The power of the lasers supported excitation times of 50 ms in each wavelength and the AOTFs allowed minimum delay ( $\sim$ 1 ms) between two colours (for example, delay between green-red for each time point), which was an important factor to assess the co-localization between markers.

Automated unbiased identification of all AP2 fluorescent spots and quantitative tracking of their dynamics as a function of time were performed as in ref. 49. Endophilin-positive assembly (EPA) budding rates represent the number of new EPAs per  $5 \times 10^4 \mu\text{m}^2$  per second, normalized to the average of the control captures. Co-localization scorings were performed on at least 100 EPAs at the plasma membrane, from three cells imaged during independent experiments.

**Signal transduction and immunoblotting.** Cells grown on 6-well plates were transfected with siRNAs and transferred to 24-well plates on the following day. Three days after transfection cells were serum starved for 4 h and stimulated at 37 °C with 5 or 100 ng ml<sup>-1</sup> EGF diluted in serum-free  $\alpha$ -MEM (Life Technologies) for the indicated times. After each time points, cells were quickly collected in Laemmli sample buffer, sonicated and boiled. Samples were run on NuPage 4–12% Bis-Tris gel (Life Technologies), transferred to nitrocellulose membranes, blocked in 5% skimmed milk and incubated successively with primary and secondary-HRP coupled antibodies and finally visualized with ECL Plus, Pico or Femto (Thermo) reagents depending on the strength of the signals. Signals were captured on a BioRad Geldoc XR+ system and analysed using the ImageLab software (BioRad). Signals used for quantifications were captured at a pre-saturation intensity. Signals were normalized to the time point zero or 30 of the control sample for the 5 ng ml<sup>-1</sup> and 100 ng ml<sup>-1</sup> experiments, respectively.

cAMP was measured using the Cyclic AMP XP chemiluminescent assay kit (Cell Signaling Technologies, 8019) following the manufacturer's instructions. Signals were measured on a FLUOstar Optima plate reader (BMG).

PathScan Intracellular Signalling and EGFR Signalling Antibody Arrays (Cell Signaling Technology 7744 and 12785, respectively) were used according to the manufacturer and fluorescent signals were measured using an Odyssey (LI-COR) reader.

**Plasma membrane isolation.** Cell surface protein isolation, excluding intracellular proteins, was performed by selective biotinylation using a cell-impermeable, cleavable biotinylation reagent (Sulfo-NHS-SS-Biotin) followed by purification of

surface proteins by affinity-purified NeutrAvidin Agarose Resin (Thermo Scientific 89881).

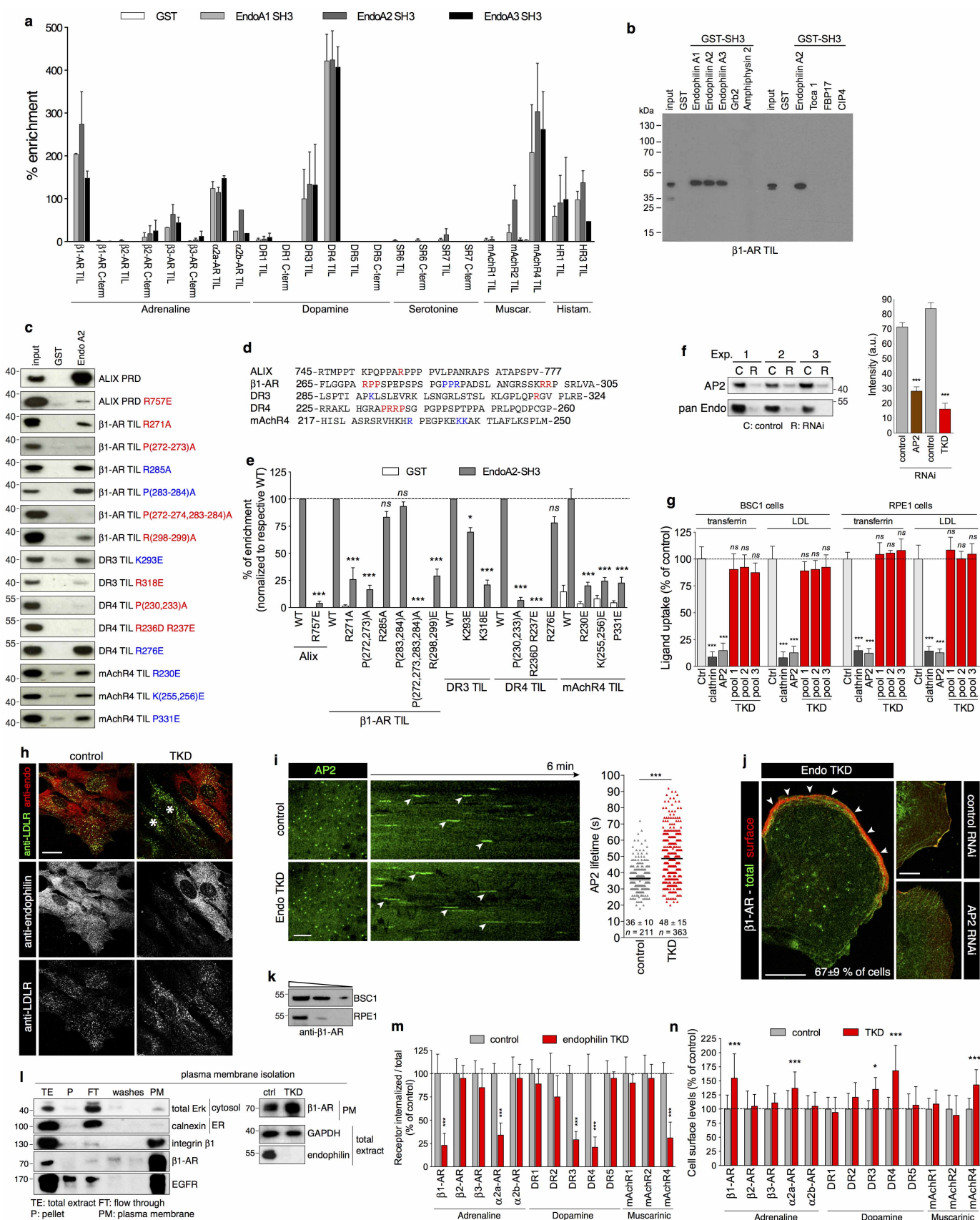
**Wound-healing cell migration assay.** BSC1 and RPE1 cells were grown in 6-well plates in their respective media for 24 h. A mechanical wound was introduced in the middle of the confluent layers of cells and the cells were allowed to grow in fresh media supplemented or not with the indicated inhibitors at 37 °C. Images were taken immediately after the wound and regularly until the end of the assay (16 h). Cell migration was measured by the reduction of free space between the two pictures and normalized to the mean of control cells.

**PC12 neurite outgrowth assay.** PC12 cells were grown in in RPMI 1640 containing 7.5% heat-inactivated horse serum, 7.5% FBS and subcultured using trypsin in type IV collagen coated flasks (BD Biocoat, Cat No 354523). Cells were detached, re-suspended in PBS, counted and electroporated with 20 pmol siRNA in 10  $\mu$ l tips using a Neon electroporation system (Life Technologies) set to a single pulse of 1,410 V for 30 ms. Cells were immediately transferred to warmed medium containing or not NGF (2.5S subunit, 100 ng ml<sup>-1</sup>) and plated onto dishes coated with poly-L-lysine/type I collagen 25:1 at a density of  $1 \times 10^4$  cells cm<sup>-2</sup>. After 2 days, cells were re-transfected with silencing RNAs using Lipofectamine RNAi MAX according to manufacturer's instructions. Cells were re-fed with or without NGF every 48 h.

**Statistical analysis.** Results shown are mean  $\pm$  standard error of the mean (s.e.m.). Statistical testing was performed using Prism 6 (GraphPad Software). Data were tested for Gaussian distribution with Kolmogorov–Smirnov test with the Dallal–Wilkinson–Lillie for corrected *P* value. In case of Gaussian distribution, the following parametric tests were used: Student's *t*-test (2 groups) or one-way ANOVA and Dunnett's test (2+ groups), as appropriated. In case of non-Gaussian distribution, the following non-parametric tests were used: two-tailed Mann–Whitney *U*-test (2 groups) or Kruskal–Wallis one-way ANOVA and Dunnett's test (2+ groups), as appropriated.

Significance of mean comparison is annotated as follow: NS, not significant; \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001. No statistical methods were used to pre-determine sample size.

45. Shaner, N. C. *et al.* Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nature Methods* **5**, 545–551 (2008).
46. Usami, Y., Popov, S. & Gottlinger, H. G. Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. *J. Virol.* **81**, 6614–6622 (2007).
47. Ford, M. G. *et al.* Simultaneous binding of PtdIns(4,5)P<sub>2</sub> and clathrin by AP180 in the nucleation of clathrin lattices on membranes. *Science* **291**, 1051–1055 (2001).
48. Marks, B. *et al.* GTPase activity of dynamin and resulting conformation change are essential for endocytosis. *Nature* **410**, 231–235 (2001).
49. Henne, W. M. *et al.* FCHO proteins are nucleators of clathrin-mediated endocytosis. *Science* **328**, 1281–1284 (2010).
50. Motley, A., Bright, N. A., Seaman, M. N. & Robinson, M. S. Clathrin-mediated endocytosis in AP-2-depleted cells. *J. Cell Biol.* **162**, 909–918 (2003).
51. Subtil, A., Hemar, A. & Dautry-Varsat, A. Rapid endocytosis of interleukin 2 receptors when clathrin-coated pit endocytosis is inhibited. *J. Cell Sci.* **107**, 3461–3468 (1994).
52. Pelish, H. E. *et al.* Secramine inhibits Cdc42-dependent functions in cells and Cdc42 activation in vitro. *Nature Chem. Biol.* **2**, 39–46 (2006).
53. Macia, E. *et al.* Dynasore, a cell-permeable inhibitor of dynamin. *Dev. Cell* **10**, 839–850 (2006).
54. Howes, M. T. *et al.* Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells. *J. Cell Biol.* **190**, 675–691 (2010).
55. Hill, T. A. *et al.* Inhibition of dynamin mediated endocytosis by the dynoles–synthesis and functional activity of a family of indoles. *J. Med. Chem.* **52**, 3762–3773 (2009).
56. Joshi, S. *et al.* The dynamin inhibitors MiTMAB and OcTMAB induce cytokinesis failure and inhibit cell proliferation in human cancer cells. *Mol. Cancer Ther.* **9**, 1995–2006 (2010).
57. Schlegel, R., Dickson, R. B., Willingham, M. C. & Pastan, I. H. Amantadine and dansylcadaverine inhibit vesicular stomatitis virus uptake and receptor-mediated endocytosis of alpha 2-macroglobulin. *Proc. Natl Acad. Sci. USA* **79**, 2291–2295 (1982).
58. Wang, L. H., Rothberg, K. G. & Anderson, R. G. Mis-assembly of clathrin lattices on endosomes reveals a regulatory switch for coated pit formation. *J. Cell Biol.* **123**, 1107–1117 (1993).
59. Gibson, A. E., Noel, R. J., Herlihy, J. T. & Ward, W. F. Phenylarsine oxide inhibition of endocytosis: effects on asialofetuin internalization. *Am. J. Physiol.* **257**, C182–C184 (1989).
60. Heuser, J. E. & Anderson, R. G. Hypertonic media inhibit receptor-mediated endocytosis by blocking clathrin-coated pit formation. *J. Cell Biol.* **108**, 389–400 (1989).
61. Larkin, J. M., Brown, M. S., Goldstein, J. L. & Anderson, R. G. Depletion of intracellular potassium arrests coated pit formation and receptor-mediated endocytosis in fibroblasts. *Cell* **33**, 273–285 (1983).
62. Yip, S. C. *et al.* Quantification of PtdIns(3,4,5)P<sub>3</sub> dynamics in EGF-stimulated carcinoma cells: a comparison of PH-domain-mediated methods with immunological methods. *Biochem. J.* **411**, 441–448 (2008).

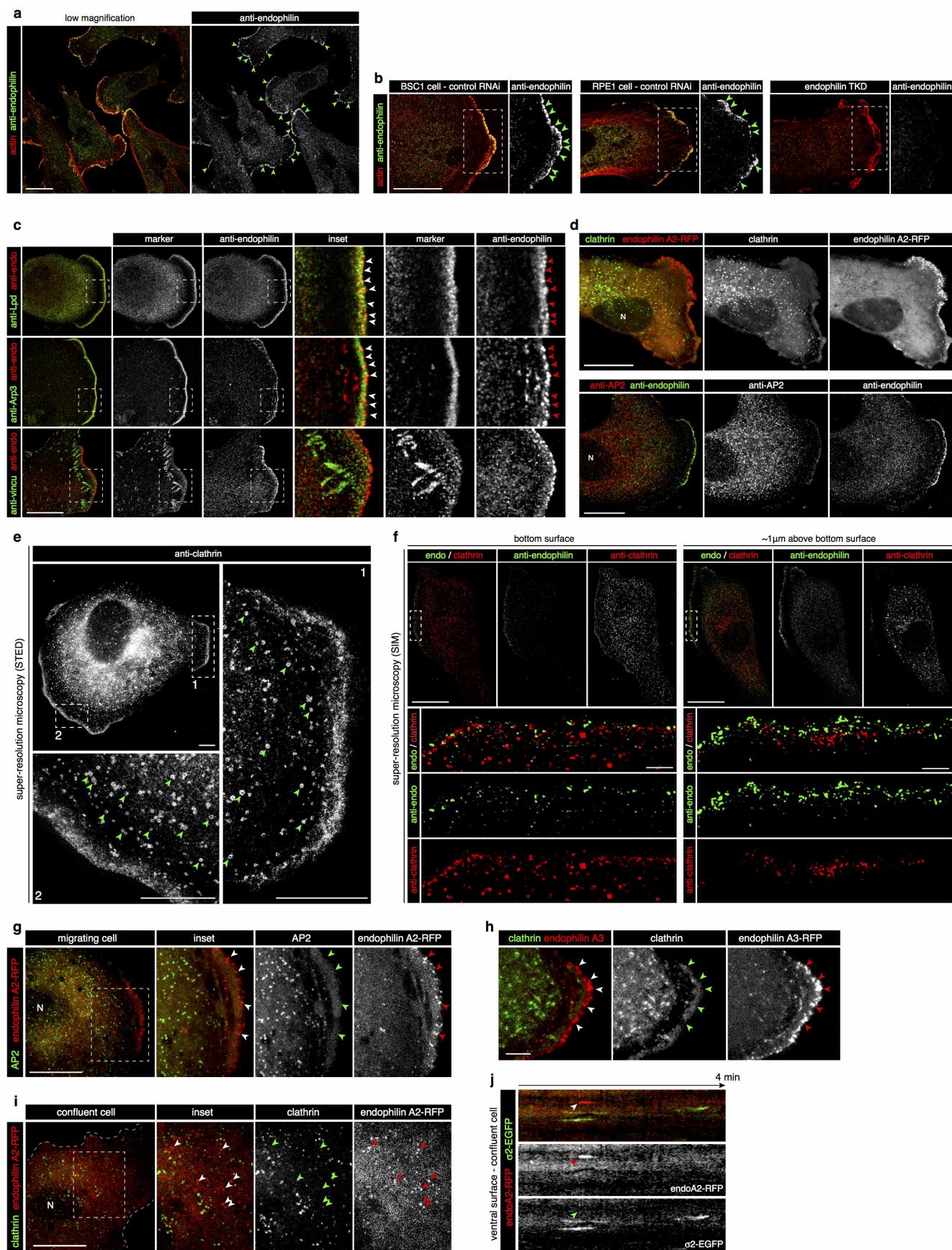


### Extended Data Figure 1 | Endophilin binds to some GPCRs and is dispensable for clathrin-mediated endocytosis. Related to Fig. 1.

**a**, Quantification of the binding of indicated cytosolic fragments of adrenergic, dopaminergic, serotonergic, muscarinic (Muscar.) and histaminergic (Histam.) receptors (see Fig. 1a for an example of one set of immunoblots) to the indicated GST-SH3 domains (or GST as control). The percentage enrichment of receptor tails in pulled-down fractions with respect to their respective levels in the original cell extracts (input) is shown (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **b**, Binding of GST-SH3 domains from BAR and F-BAR containing proteins to  $\beta_1$ -AR third intracellular loop (TIL), showing specificity for endophilin-A SH3 domains. **c**, Pull-down experiments using GST-SH3 domains of endophilin-A2 (or GST as a control) with EGFP-tagged fragments coding for the TILs of the indicated receptors bearing the depicted mutations. Mutations decreasing binding are shown in red; the ones without significant effects are in blue. ALIX PRD domain and its R757E mutant<sup>58</sup> were used as positive and negative controls. **d**, Amino acid sequence around point mutations tested showing potential endophilin-binding motifs. **e**, Quantification of the binding of indicated mutated cytosolic fragments (see **c** for an example of one set of immunoblots) to the indicated GST-SH3 domains (or GST as control). The percentage enrichment of receptors tails in pulled-down fractions with respect to their respective levels in the original cell extracts (input) is shown (means  $\pm$  s.e.m. from three independent experiments; NS, non-significant,  $*P < 0.05$ ,  $***P < 0.001$ , one-way ANOVA and Dunnett's test versus respective wild-type). **f**, Verification of AP2 and Endo1+2+3 depletion by RNAi from three independent control ('C') and RNAi ('R') experiments. Immunoblot (left) and bar graph (right) representations of the data are shown (means  $\pm$  s.e.m. from three independent experiments;  $***P < 0.001$ , Student's *t*-test versus respective control). **g**, Transferrin and LDL uptake was inhibited by clathrin and AP2 RNAi but not by control (Ctrl) or endophilin-A triple RNAi (TKD, three different pools). Ligand uptake was measured by flow cytometry (mean  $\pm$  s.e.m., 10,000 cells per experiment,  $n = 3$  independent experiments; NS, non-significant,  $***P < 0.001$ , one-way ANOVA and Dunnett's test versus respective control). **h**, Anti-LDLR antibody uptake assay in control or TKD RPE1 cells. Cells were immunostained for anti-LDLR (post feeding) and endophilin (red). Cells strongly depleted for endophilin (\*) showed similar LDLR endocytosis to control cells, showing that LDLR endocytosis is not dependent on endophilin-A proteins (images representative of ten captures). **i**, Dynamics of clathrin-coated pits in BSC1 cells marked by  $\sigma$ 2-EGFP (AP2) and acquired by live-cell

spinning-disk confocal microscopy imaging. Left pictures show representative images from the bottom surface of cells; right pictures are kymographs (time projections) representative of five captures from control and Endo1+2+3 RNAi cells. Right: plot of individual lifetimes of clathrin/AP2 pits calculated from live imaging of 4 cells for each condition. Those had significant longer lifetimes in TKD cells, perhaps related to a reduced synaptojanin recruitment to membranes and thus an increase in PtdIns(4,5)P<sub>2</sub> levels<sup>5,20</sup> (bars, means;  $n = 211$  and 363 from 3 independent experiments, mean  $\pm$  s.e.m. is written at the bottom,  $***P < 0.001$ , Mann-Whitney *U*-test). **j**, Surface staining of  $\beta_1$ -AR using the HA- $\beta_1$ -AR-EGFP construct. Control, AP2 RNAi and endophilin TKD cells were directly stained for steady-state accumulation of  $\beta_1$ -AR at their cell surface and imaged using a confocal microscope (images are representative of five captures). In endophilin TKD cells  $\beta_1$ -AR accumulated in a limited area of the plasma membrane, similar to the area stained with lamellipodin in Extended Data Fig. 2c. This panel complements that in Fig. 1b showing a lack of internalization of  $\beta_1$ -AR in TKD cells. **k**, Immunoblots showing the expression of  $\beta_1$  adrenergic receptor ( $\beta_1$ -AR) in BSC1 and RPE1 cells. Decreasing amount of cell extracts were loaded from left to right. **l**, Left: isolation of a plasma membrane fraction (see Methods) showed accumulation of  $\beta_1$ -adrenergic receptors after endophilin RNAi. Erk was used to monitor cytosolic proteins, calnexin endoplasmic reticulum proteins and integrin  $\beta_1$ ,  $\beta_1$ -adrenergic receptor ( $\beta_1$ -AR) and EGFR to monitor plasma membrane proteins. Note that the plasma membrane fraction (PM) was enriched in plasma membrane proteins and did not contain cytosolic ERK or ER-localized calnexin. Right: levels of  $\beta_1$ -AR were enriched in the plasma membrane fraction from TKD cells versus control cells. **m**, Bar graph shows the amounts of internalized/total depicted HA-receptor-EGFP in control (grey) or endophilin TKD (red) samples (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments;  $***P < 0.001$ , Student's *t*-test versus respective control). **n**, Surface staining of HA-receptor-EGFP constructs (see Fig. 1b, c). Cells were directly stained with anti-HA antibody at 4 °C and measured using a plate reader to assay the steady-state accumulation of receptors at the cell surface. The cell surface signals (HA) were corrected for expression levels (EGFP) and the values were normalized to the mean of the control cells (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments;  $***P < 0.001$ , Student's *t*-test versus respective control). The same receptors that showed a decreased endocytosis (**m**) show an increased surface accumulation here, an effect that is dependent on ligand activation. Scale bars: 20  $\mu$ m (**h**), 10  $\mu$ m (**j**) and 5  $\mu$ m (**i**).



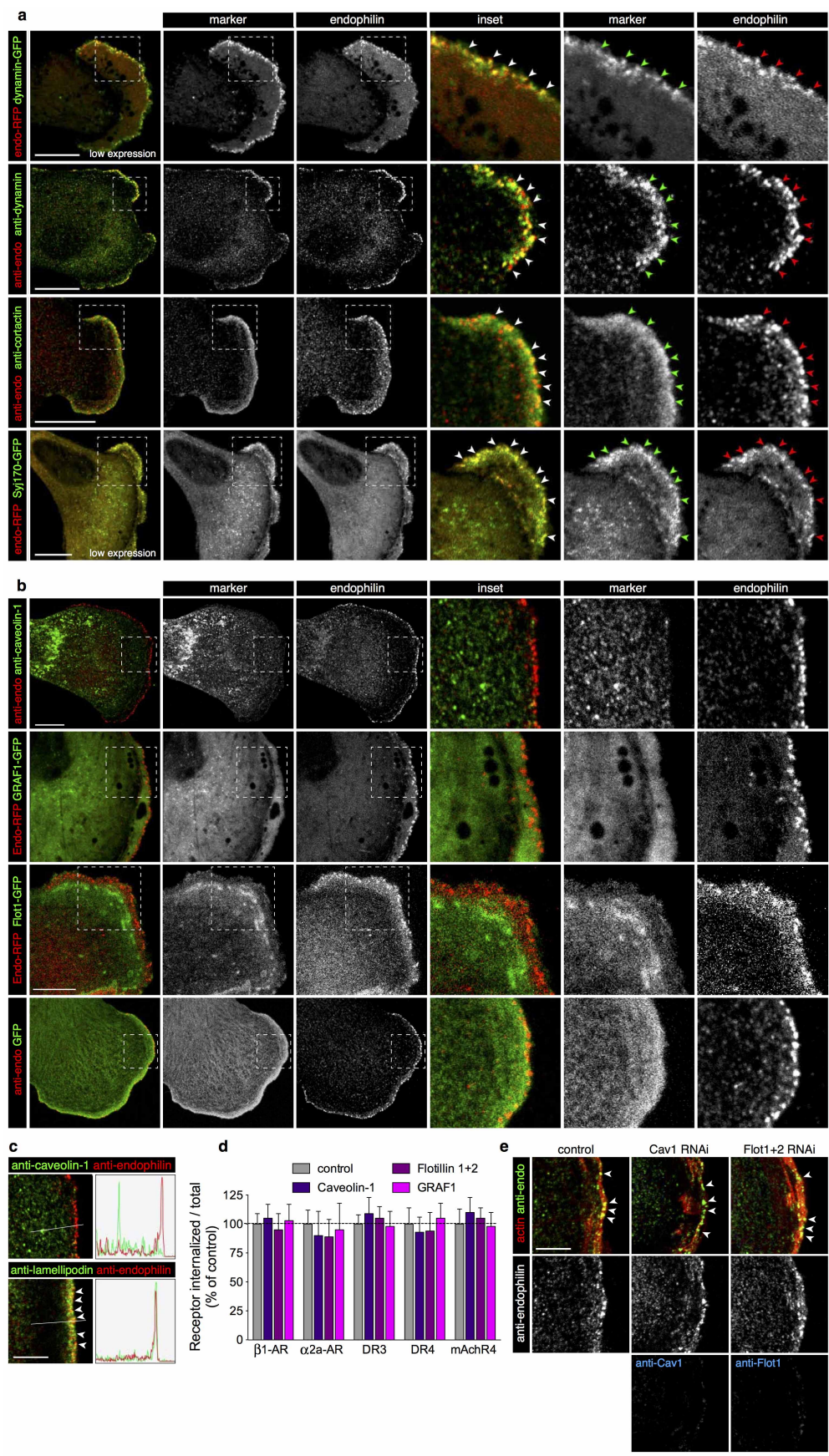


**Extended Data Figure 2 | Endophilin and clathrin did not co-localize at the leading edge of cells by standard and super-resolution microscopy.** Related to Fig. 1.

**a, b**, Confocal microscopy images (optical planes from the middle of cells) showing endogenous staining of endophilin (green) in BSC1 cells (low magnification in **a** and higher magnification in **b**, left panel) and in a normal diploid RPE1 cell (middle). Images in **b** were oriented with leading edges (identified by phalloidin staining (actin, red)) pointing to the right. Arrowheads point to endophilin puncta on the leading edge of these cells. Right: there was no anti-endophilin staining in endophilin-A TKD cells, validating the specificity of the antibody. Insets are zooms of the boxed areas. **c**, Confocal microscopy images showing endogenous staining of lamellipodin, Arp3 or vinculin (all green) and endophilin (red) in BSC1 cells. Insets are zooms of the boxed areas. Endophilin co-localized with lamellipodin and Arp3 at the leading edge but not with vinculin (a marker of focal adhesions). **d**, Top: same data as in Fig. 1c but showing the individual channels. Confocal microscopy images showing clathrin (EGFP-LCa, green) and endophilin (endophilin A2-RFP, red) in a live BSC1 cell. 'N' denotes the nucleus. Bottom: confocal microscopy images showing endogenous staining of  $\alpha$ -adaptin (AP2), and endophilin (red) in BSC1. Endogenous staining also shows an enrichment of endophilin at the leading edge. **e**, Super-resolution stimulated emission depletion (STED) microscopy images of a BSC1 cell immunostained for clathrin and showing coated pits and vesicles labelled at the ventral surface of the cell but not at the leading edges where the staining was diffused.

**f**, Super-resolution structured illumination microscopy (SIM) images of a BSC1 cell immunostained for clathrin (red) and endophilin (green). Note the absence of co-localization between the two markers. **g**, Confocal microscopy images (bottom surface optical section) showing the localization of AP2 ( $\sigma$ 2-EGFP, green) and endophilin (EndoA2-RFP, red) in a live migrating cell. Arrowheads point to endophilin A2 puncta at the leading edge of the cell; these are negative for AP2. Inset is a zoom of the boxed area. **h**, Confocal microscopy images showing a lack of co-localization of clathrin (EGFP-LCa, green) and the other endophilin-A expressed in BSC1 cells: endophilin A3 (endophilin A3-RFP, red). Arrowheads point to endophilin A3 puncta at the leading edge of live BSC1 cells; these are negative for clathrin puncta. **i**, Confocal microscopy images (optical section of bottom surface) showing clathrin (EGFP-LCa, green) and endophilin (EndoA2-RFP, red) localization in a live confluent BSC1 cell (thus no leading edge). Neighbouring cells were not transfected. Arrowheads point to endophilin A2 puncta at the bottom surface of the cell; most of these are negative for AP2. Inset is a zoom of the boxed area. **j**, Representative kymographs of AP2 ( $\sigma$ 2-EGFP) and endophilin (EndoA2-RFP) from live-cell imaging acquired at the ventral surface of a confluent cell. The arrowhead points to an endophilin-positive track devoid of AP2. Images and kymographs are representative of at least ten captures, from three independent experiments. Scale bars: 20  $\mu$ m (**a, d**), 10  $\mu$ m (**b, c, f, e**), 5  $\mu$ m (**e, h**) and 1  $\mu$ m (**f** insets).

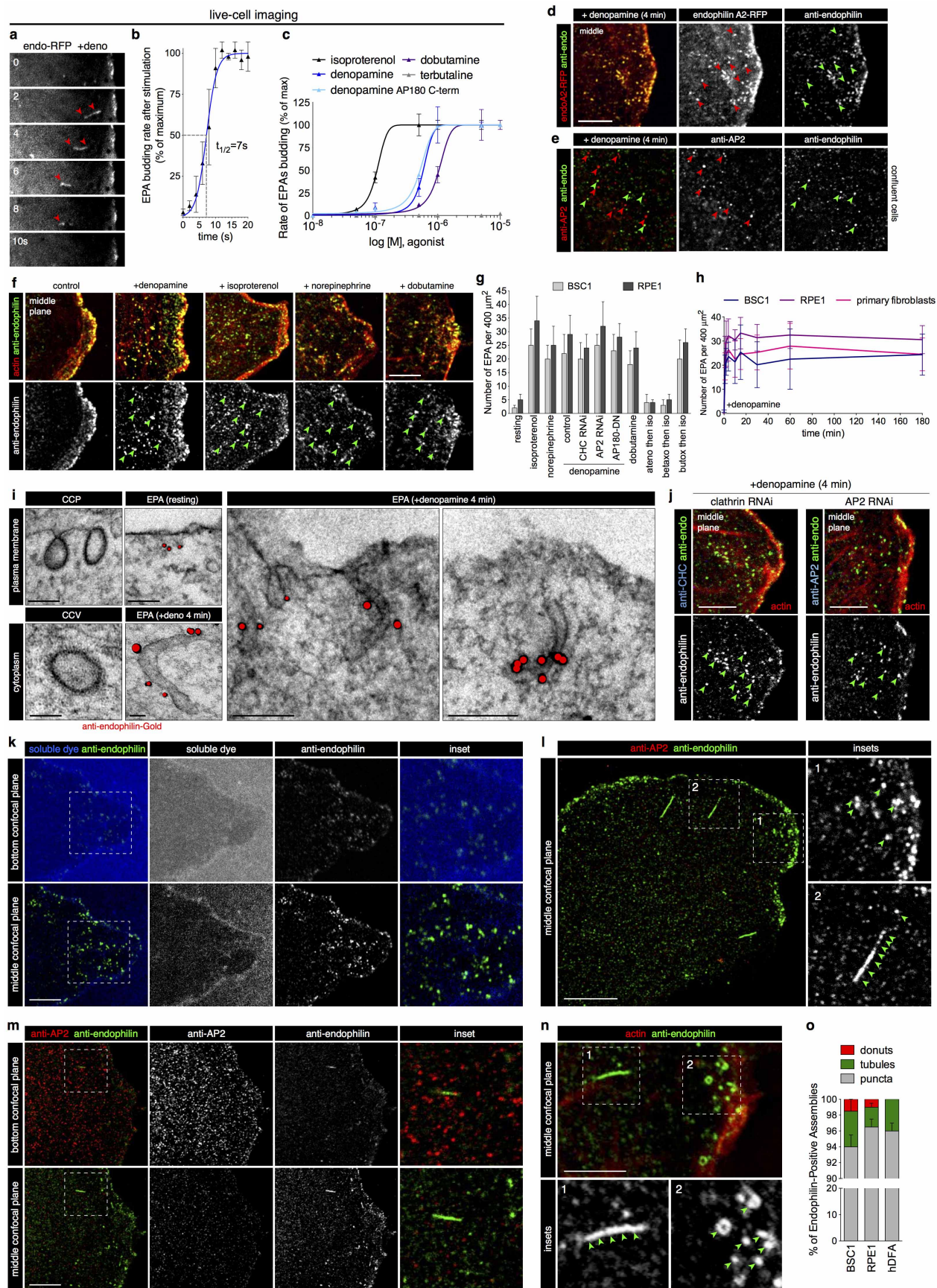






**Extended Data Figure 3 | Endophilin co-localized with dynamin, cortactin and synaptojanin but not with known clathrin-independent endocytosis markers.** Related to Fig. 1. **a**, Representative confocal images showing the co-localization of endophilin (endoA2-RFP or anti-endophilin, red) with dynamin (dynamin2-EGFP expressed at low levels, or anti-dynamin for endogenous protein (green)), cortactin (anti-cortactin for endogenous protein) or synaptojanin (synaptojanin 1-170-EGFP expressed at low levels). Insets are zooms of the boxed areas. **b**, Representative confocal images showing the absence of co-localization between endophilin (anti-endophilin, or endoA2-RFP, red) and caveolin 1 (endogenous), GRAF1-EGFP and flotillin 1-EGFP. The bottom row shows the background co-localization of endophilin with a soluble protein (EGFP). Arrowheads point to endophilin puncta at the leading edge of cells; these were negative for the markers tested. Insets are

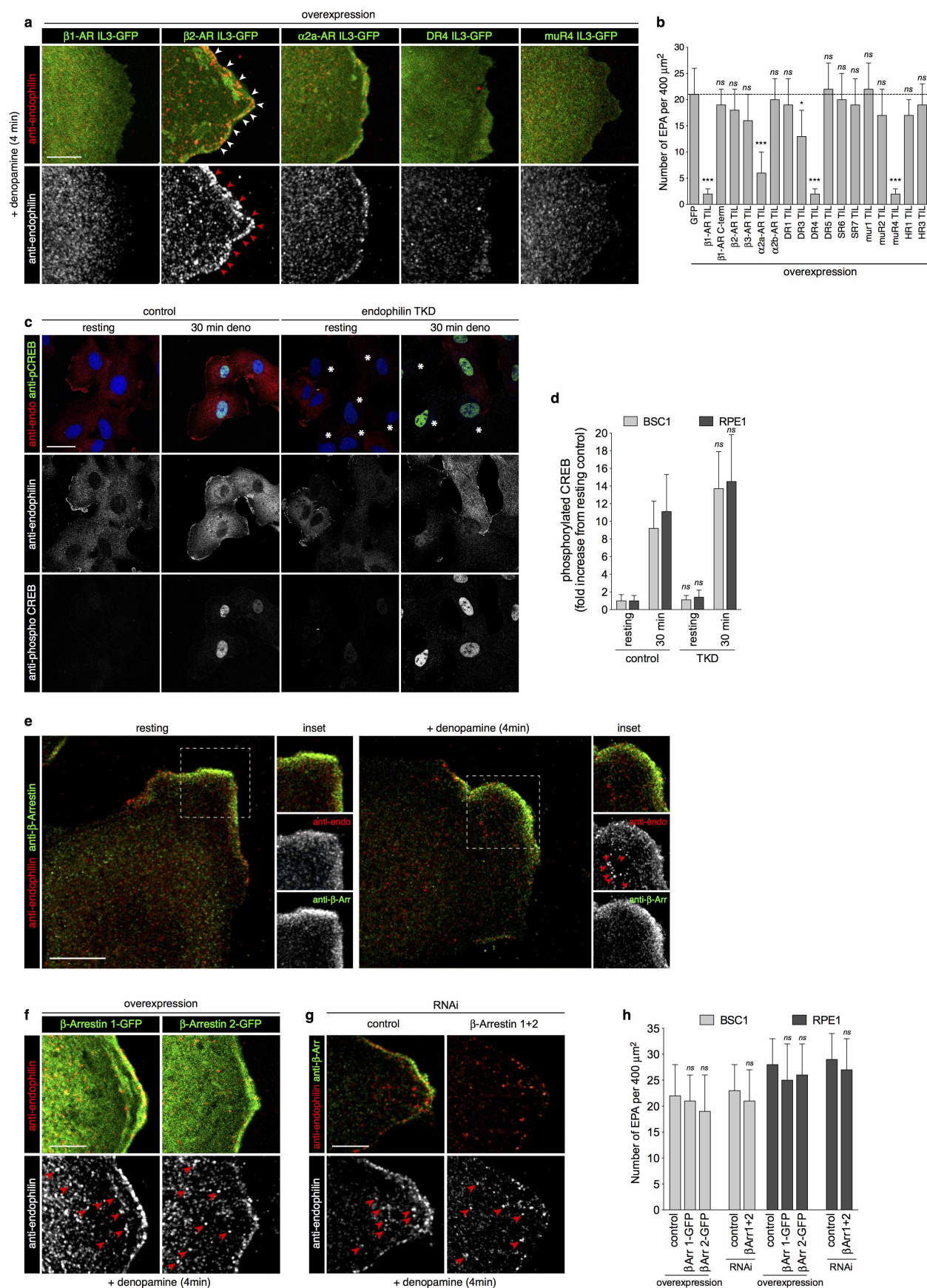
magnifications of the boxed areas. **c**, Co-localization of endogenous endophilin (red) with lamellipodin, but not caveolin-1 (green). Images were oriented with the leading edges to the right. Arrowheads point to co-localization between markers. Intensity profiles were acquired along the indicated lines. **d**, Receptor uptake assay (HA-receptor-EGFP internalized corrected for their total levels) in cells pre-treated with the indicated RNAi (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **e**, Confocal images of cells treated with caveolin-1 (Cav1) or flotillin 1 and 2 (Flot1+2) or control siRNA and immunostained for endophilin (green) and actin (red). Cells were counter-stained (blue) for caveolin 1 (Cav1) or flotillin-1 (Flot1) to ascertain for the depletion of the targeted proteins in the cells imaged. Images are representative of at least ten captures, from three independent experiments. Scale bars: 10  $\mu$ m (**a**) and 5  $\mu$ m (**b**, **c** and **e**).



**Extended Data Figure 4 | Formation of endophilin-positive tubules and vesicles upon  $\beta_1$ -adrenergic receptor activation.** Related to Fig. 1. **a**, Images from a time-lapse acquired by spinning-disk confocal microscopy of a BSC1 cell expressing low levels of endophilin-A2-RFP, after addition of 10  $\mu$ M of denopamine ( $t = 0$ ). Arrowheads point to a tubulo-vesicular carrier formed upon stimulation and moving from the cell edge (right) to the cell centre (left). See also Supplementary Video 3. **b**, The budding rate of endophilin-positive assemblies (EPAs), measured as number of EPAs  $\mu\text{m}^{-2}\text{s}^{-1}$ , after stimulation with denopamine (added at  $t = 0$ ). Data were acquired from data sets similar to Supplementary Video 3 and normalized to the maximum budding rate (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **c**, Budding of endophilin-mRFP-positive structures determined by live-cell imaging after addition of various  $\beta_1$ -AR (isoproterenol, dobutamine and denopamine) or  $\beta_2$ -AR (isoproterenol, terbutaline) agonists concentrations. Data are expressed as the percentage of maximum budding. See also Supplementary Video 3. Terbutaline, the specific  $\beta_2$ -AR agonist, does not stimulate EPA production. Inhibition of clathrin-mediated endocytosis by AP180 C terminus expression does not affect EPA formation stimulated by the  $\beta_1$ -AR agonist denopamine (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **d**, Representative confocal images (optical planes located at the middle of cells) showing the co-localization of endophilin A2-RFP (red) with endogenous endophilin (green) in a BSC1 cell stimulated with denopamine (10  $\mu$ M for 4 min), validating the targeting of exogenously expressed endophilin in our experiments. **e**, Representative confocal images (optical section located at the middle of cells) of confluent cells stimulated for 4 min with 10  $\mu$ M denopamine and stained for endogenous endophilin (green) and  $\alpha$ -adaptin (AP2, red). **f**, Representative confocal images (optical section located at the middle of cells) of sparse (**f**) or confluent (**e**) cells stimulated for 4 min with the indicated adrenergic agonists and stained for endogenous endophilin (green) and actin (phalloidin, red). **g**, Quantification of EPAs in BSC1 and RPE1 cells upon stimulation with adrenergic receptor agonists in control, AP2 RNAi or AP180 dominant-negative-expressing cells or in cells pre-treated (5 min) with beta-blockers before stimulation (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **h**, Quantification of the number of endogenous EPAs per 400  $\mu\text{m}^2$  in BSC1, RPE1 cells and in primary fibroblasts upon stimulation with 10  $\mu$ M

denopamine for the indicated amounts of time (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **i**, Immuno-electron microscopy of BSC1 cells fixed 4 min after addition of 10  $\mu$ M denopamine (deno) to the medium or in absence of stimulation (control). Anti-endophilin antibodies were detected with a gold-conjugated secondary antibody (red balls). The two pictures on the left show clathrin-coated pits (top) and a potential vesicle (bottom). The other images are membrane structures at the plasma membrane and in the cytoplasm positive for the anti-endophilin antibody (endogenous endophilin). **j**, Representative confocal images (optical section located at the middle of cells) of cells depleted for clathrin (left, same image as in Fig. 1e) or AP2 and stimulated with denopamine (10  $\mu$ M) for 4 min before fixation. Endogenous endophilin (green) clathrin or AP2 (blue) and actin (phalloidin, red) were immunostained. **k**, Confocal microscopy images of a stimulated (10  $\mu$ M denopamine for 4 min) RPE1 cell grown on a poly-lysine-coated coverslip, incubated in soluble lysine-fixable Alexa647 dye, fixed and immunostained for endophilin (green). On the confocal plane taken at the bottom of the cell (dye in focus), no endophilin spots can be seen (out of focus). On a confocal plane at the middle of the cell (dye outlines the cell cross-section) EPAs are in focus, indicating that they are in the cytoplasm and not at the cell surface. **l**, Example of a BSC1 cell treated and immunostained as in **m** and imaged at the middle of the cell and showing several endophilin-positive tubules and diffraction-limited puncta. **m**, Confocal microscopy images of a stimulated (10  $\mu$ M denopamine for 4 min) BSC1 cell fixed and immunostained for  $\alpha$ -adaptin (AP2, red) and endophilin (green). On the confocal plane taken at the bottom of the cell AP2 spots (known to be located at the plasma membrane) are in focus but EPAs are out of focus. On a confocal plane at the middle of the cell AP2 spots are out of focus but EPAs are in focus, indicating that they are inside the cytoplasm and not at the plasma membrane. **n**, Confocal microscopy images of EPAs labelled by endophilin. Arrowheads show diffraction-limited punctates, tubules and 'doughnut-like' vesicles structures. **o**, Quantification of the occurrence in stimulated BSC1, RPE1 and human primary dermal fibroblasts of the three different endophilin-positive structures: diffraction-limited punctates (grey), tubules (green) and 'doughnut-like' vesicles (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). Images are representative of at least ten captures, from three independent experiments. Scale bars: 5  $\mu\text{m}$  apart from **i**, which is 250 nm.

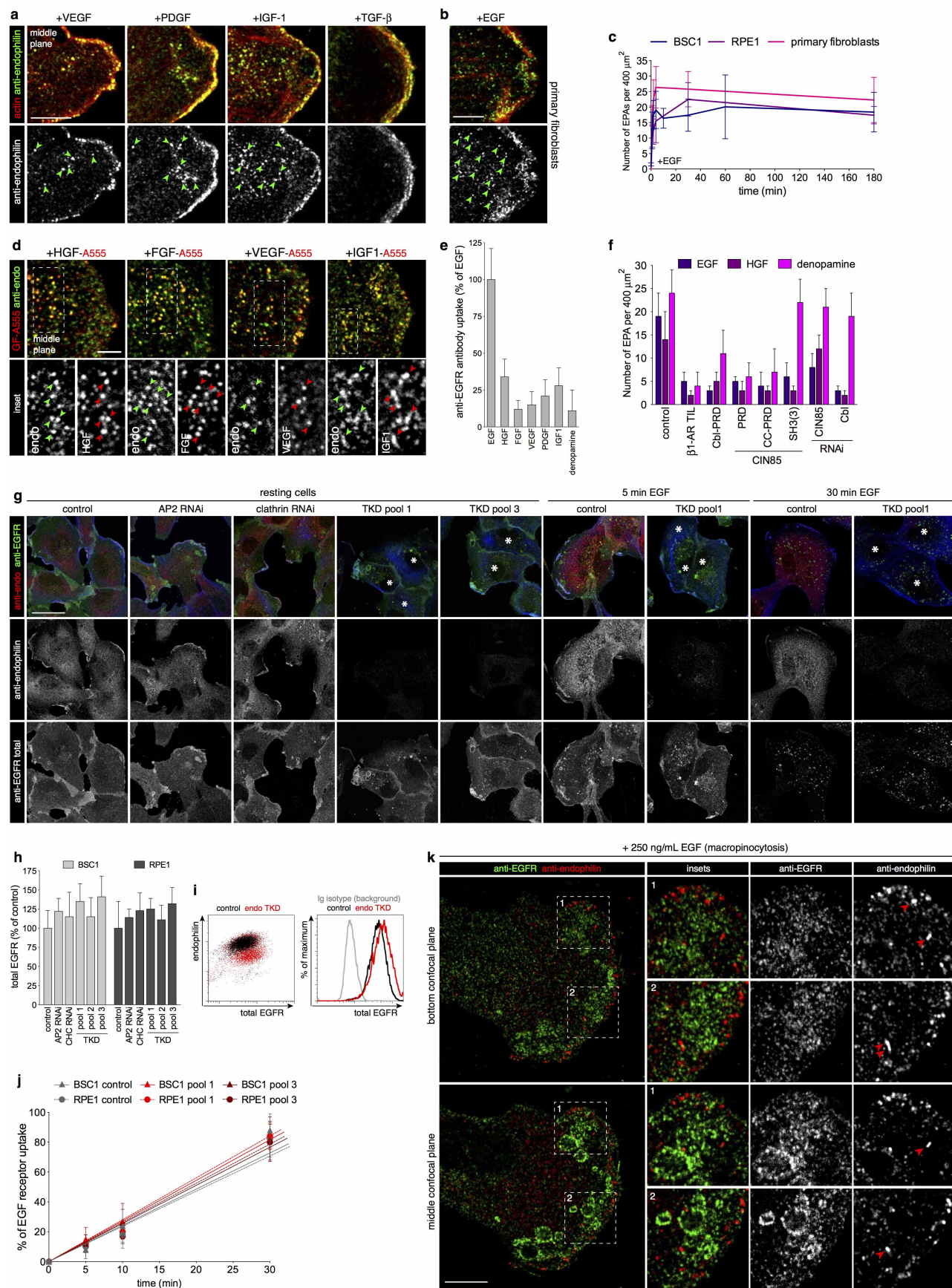




### Extended Data Figure 5 | $\beta 1$ and $\beta 2$ arrestins were not essential in $\beta 1$ -adrenergic receptor agonist stimulation of endophilin vesicle formation.

Related to Fig. 1. **a**, Representative confocal images (optical planes located at the middle of cells) of cells overexpressing the indicated EGFP-tagged receptors TIL (green), stimulated with 10  $\mu$ M denopamine for 4 min and immunostained for endophilin (red) (images are representative of ten captures from three independent experiments). **b**, Quantification of the number of endogenous EPAs in cells overexpressing the indicated EGFP-tagged receptors TIL and stimulated with 10  $\mu$ M denopamine for 4 min (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant,  $*P < 0.05$ ,  $***P < 0.001$ , one-way ANOVA and Dunnett's test versus GFP). **c**, Confocal microscopy images of control and endophilin TKD cells, resting or stimulated (10  $\mu$ M denopamine for 4 min) and immunostained for phosphorylated CREB (pCREB, green), endophilin (red) and DNA (DRAQ5, blue) (images are representative of ten captures from three independent experiments). Asterisks

denote cells with strongly reduced endophilin levels. **d**, Plate reader quantification of BSC1 or RPE1 cells treated as in **a** (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant, Student's *t*-test versus respective control). **e**, Confocal microscopy images of control and stimulated (10  $\mu$ M denopamine for 4 min) cells immunostained for  $\beta$ -arrestin (green) and endophilin (red). EPAs (arrowheads) did not contain  $\beta$ -arrestin (images representatives of ten captures from three independent experiments). **f**, **g**, Confocal microscopy images of cells overexpressing  $\beta$ -arrestin-1 or 2-EGFP (**f**, green) or  $\beta$ -arrestin 1 + 2 KD cells (**g**), stimulated with denopamine (10  $\mu$ M for 4 min), fixed and immunostained for endophilin (red) and  $\beta$ -arrestin (**g**, green) (images representatives of at least ten captures). **h**, Plate reader quantification of BSC1 or RPE1 cells treated as in **e–g** (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant, one-way ANOVA and Dunnett's test versus respective controls). Scale bars: 40  $\mu$ m (**c**) and 5  $\mu$ m (**a**, **e–g**).





**Extended Data Figure 6 | Several growth factors stimulated the formation of endophilin-positive assemblies.** Related to Fig. 2. **a**, Confocal images

(optical planes from the middle of cells) of BSC1 cells treated for 4 min with the indicated growth factors (all at  $10 \text{ ng ml}^{-1}$ ) and immunostained for endogenous endophilin-A (green). Cells were oriented with their leading edges (identified with phalloidin staining (actin, red)) pointing to the right (images are representative of ten captures from three independent experiments).

Arrowheads show internal vesicles positive for endogenous endophilin in the VEGF, PDGF and IGF-1-treated but not TGF- $\beta$ -treated cells (see also quantification in Fig. 2c). **b**, EPA production in a primary fibroblast from an adult donor treated for 4 min with  $50 \text{ ng ml}^{-1}$  EGF and immunostained for endogenous endophilin-A (green). The image is representative of eight captures.

**c**, Quantification of the number of endogenous EPAs per  $400 \mu\text{m}^2$  in BSC1, RPE1 cells or primary fibroblasts upon stimulation with  $50 \text{ ng ml}^{-1}$  EGF for the indicated amounts of time (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments).

**d**, Confocal section of a BSC1 cell fixed 4 min after addition of  $50 \text{ ng ml}^{-1}$  Alexa555-labelled EGF, HGF, FGF, VEGF and IGF-1 to the medium. Endogenous endophilin (green) was detected as in **a**. Arrowheads point to internal vesicles positive for endogenous endophilin and containing internalized growth factors (images are representatives of at least eight captures).

**e**, Plate reader quantification of BSC1 cell incubated with a monoclonal antibody anti-EGFR (13A9, green), that does not interfere with EGF binding, and  $50 \text{ ng ml}^{-1}$  of the indicated ligands. There is no cross-talk between EGFR and activation of other growth factor receptors or activation of

$\beta 1$  adrenergic receptors. Thus, endophilin-positive structures only have EGFR when stimulated by EGF (see Fig. 2d, e), and likewise on  $\beta 1$  adrenergic receptor activation there is no EGFR in the endophilin-positive structures.

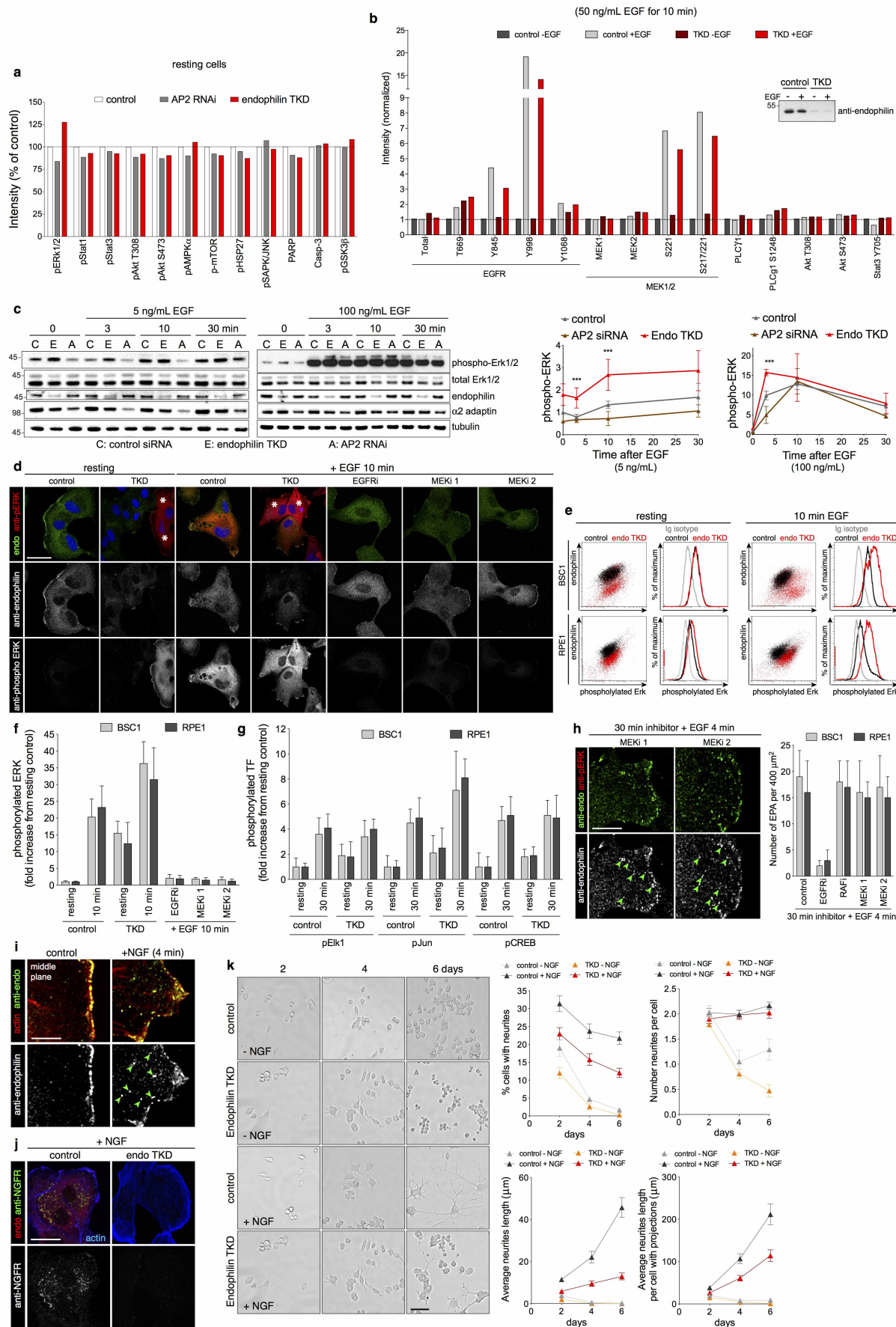
This reinforces the theme that one needs specific receptor activation to transduce a signal across the membrane to activate EPA formation (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments).

**f**, Quantification of the number of EPAs in cells depleted for CIN85 or Cbl or overexpressing  $\beta 1$ -AR TIL, Cbl-PRD or CIN85-PRD, -CC-PRD, SH3(3) and stimulated with EGF (blue), HGF (purple) or denopamine (fuchsia) (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **g**, Confocal images of control, AP2, clathrin or endophilin TKD (pools 1 and 3) treated cells, fixed and immunostained for EGFR (green) and endophilin (red) (images are representative of at least ten captures). Asterisks denote cells with strongly reduced endophilin levels.

**h**, Plate reader quantification of BSC1 and RPE1 treated as in **g** (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **i**, Flow cytometry profiles of total EGFR levels in control (black) or TKD (red) cells.  $n = 50,000$  cells for each conditions.

**j**, Plate reader quantification of control or endophilin TKD BSC1 and RPE1 treated with  $50 \text{ ng ml}^{-1}$  EGF for the indicated time (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments).

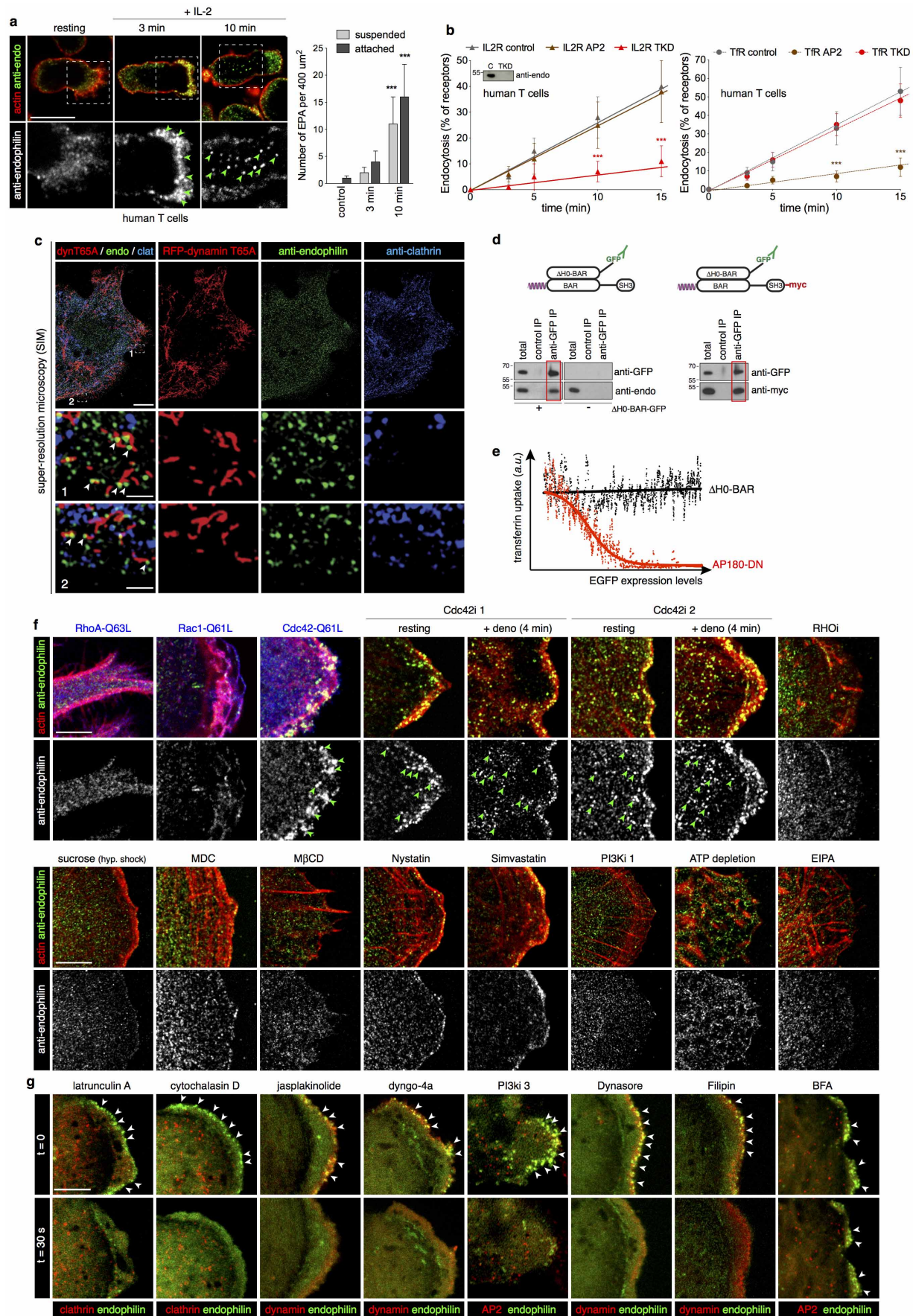
**k**, Confocal images of a BSC1 cell stimulated for 5 min with  $250 \text{ ng ml}^{-1}$  EGF, fixed and immunostained for EGFR (green) and endophilin (red) (images are representative of eight captures). Note that EPAs (arrowheads) are distinct from the large vacuole-like structures which are called macropinosomes, seen best in the middle confocal plane. Scale bars:  $40 \mu\text{m}$  (**g**),  $10 \mu\text{m}$  (**k**) and  $5 \mu\text{m}$  (**a-c**).



**Extended Data Figure 7 | MAPK signalling and neurite outgrowth in endophilin TKD cells.** Related to Fig. 2. **a**, Signals from PathScan antibody arrays reporting various intracellular signalling pathways. Intensities were normalized to the mean of the controls. Of all the signalling pathways tested endophilin TKD leads to a higher level of ERK phosphorylation (a component of the MAP kinase signalling cascade) in resting cells (mean,  $n = 2$  independent experiments). **b**, Signals from PathScan antibody arrays reporting various steps within the EGFR signalling cascade. Intensities were normalized to the mean of the controls. The inset shows the level of endophilin knockdown in TKD cells. Endophilin TKD does not prevent phosphorylation of the EGFR at Y998 and thus does not account for defective endocytosis. Endophilin TKD does not prevent MEK phosphorylation on the RAF phosphorylation site (S217/S221), important for MEK activation (mean,  $n = 2$  independent experiments). **c**, Immunoblots from control (C), endophilin TKD (E) or AP2 RNAi (A) cells stimulated with 5 (left) or 100 ng ml<sup>-1</sup> (right) EGF for the indicated times. Quantification of the signals from three independent experiments, normalized to the mean of the controls at  $t = 0$  are shown below. As in **a** the resting pERK levels are increased in endophilin RNAi cells, probably because of an accumulation of receptors on the cell surface. At low concentrations of EGF (5 ng ml<sup>-1</sup>) AP2 RNAi largely prevents MAP kinase signalling while endophilin RNAi does not (but rather an enhanced level of signalling). At higher concentrations of EGF (100 ng ml<sup>-1</sup>) there is much stronger signalling at an early time point in endophilin RNAi, pointing again to the importance of this pathway in reducing basal signalling (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; \*\*\* $P < 0.001$ , one-way ANOVA and Dunnett's test versus control  $t = 0$ ). **d**, Confocal microscopy images of resting or stimulated (50 ng ml<sup>-1</sup> EGF for 10 min) cells pre-treated with endophilin TKD (3 days) or the indicated inhibitors for 30 min before stimulation and immunostained for phosphorylated ERK1/2 (pERK, red), endophilin (green) and DNA (DRAQ5, blue) (images are representatives of at least ten captures). Asterisks denote cells with strongly reduced endophilin levels. **e**, Flow cytometry analysis of phosphorylated Erk1/2 levels in resting or stimulated (50 ng ml<sup>-1</sup> EGF for 10 min) control (black) and endophilin TKD (red) cells.  $n = 50,000$  cells for each conditions. **f**, Plate reader quantification of BSC1 or

RPE1 cells treated as in **d**. Thus by immunoblotting and by immunostaining there is an increase in pERK in resting conditions in endophilin RNAi cells. Note the stronger increase in basal and stimulated pERK levels in endophilin TKD cells (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **g**, Plate reader quantification of control and endophilin TKD BSC1 or RPE1 cells resting or stimulated (50 ng ml<sup>-1</sup> EGF for 30 min) and immunostained for phosphorylated Elk1, Jun or CREB (pElk1, pJun or pCREB) (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **h**, Left: confocal images (optical planes from the middle of cells) of BSC1 cells pre-treated with MEKi 1 or MEKi 2 for 30 min and stimulated with EGF (50 ng ml<sup>-1</sup>, 4 min) and immunostained for endogenous endophilin-A (green) (images are representatives of six captures). Endophilin-positive puncta formation in response to growth factors is not dependent on MEK activation. Right: quantification of the number of endogenous EPAs per 400  $\mu$ m<sup>2</sup> in BSC1 or RPE1 cells pre-treated for 30 min with the indicated inhibitors and stimulated with EGF (50 ng ml<sup>-1</sup>, 4 min) (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **i**, NGF stimulates endophilin-positive puncta formation in RPE1 cells. Confocal images (optical planes from the middle of cells) of an RPE1 cell treated for 4 min with NGF (10 ng ml<sup>-1</sup> for 4 min), fixed and immunostained for endogenous endophilin-A (green) (images are representatives of six captures). Cells were oriented with the leading edges (identified with phalloidin staining (actin, red)) pointing to the right. Arrowheads show internal vesicles positive for endogenous endophilin. **j**, Confocal images of cells treated with endophilin (endo TKD) or control siRNA and incubated with anti-NGFR antibodies (green) and 10 ng ml<sup>-1</sup> NGF for 10 min at 37 °C, surface stripped on ice, fixed and immunostained for endophilin (red) and actin (blue) (images are representatives of eight captures). **k**, Neurite extension assays. PC12 cells pre-treated with control or endophilin TKD were stimulated or not with 100 ng ml<sup>-1</sup> NGF for up to 6 days (images are representatives of at least 12 images). Quantifications show the percentage of cells with neurites (top left), average number of neurites per cell (top right), average neurite length (bottom left) and average neurite length per cell having neurites (bottom right) (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). Scale bars: 40  $\mu$ m (**d**), 20  $\mu$ m (**k**) and 5  $\mu$ m (**h–j**).

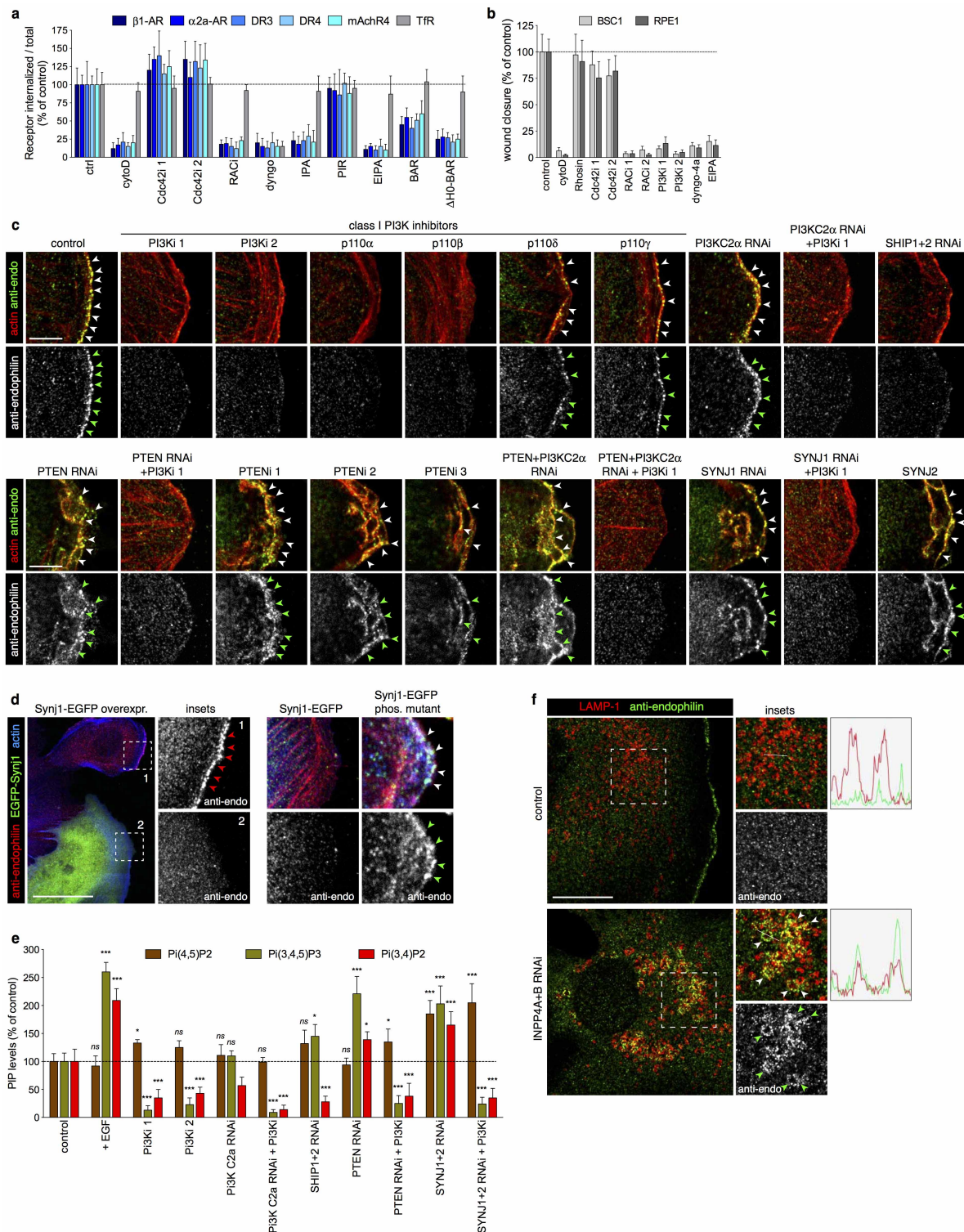




**Extended Data Figure 8 | IL-2R uptake in endophilin TKD human T cells and chemical and genetic perturbations of FEME.** Related to Figs 3 and 4. **a**, Confocal microscopy images (focal plane at the middle of cells) of human T cells immunostained for endogenous endophilin A (green) (images are representatives of at least ten images). Cells were either fixed directly ('resting') or 3 and 10 min after addition of IL-2 (500 pM). Scale bars, 5  $\mu$ m (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; \*\*\* $P < 0.001$ , Student's  $t$ -test versus respective controls). **b**, Internalization rates of IL-2R and TfR in control, AP2 or endophilin TKD human T cells, measured by a plate reader (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; \*\*\* $P < 0.001$ , Student's  $t$ -test versus respective control time points). Inset: immunoblots of cell extracts used validated endophilin TKD. **c**, Super-resolution structured illumination microscopy (SIM) images of BSC1 cells overexpressing RFP-dynamin1-T65A mutant and immunostained for endophilin (green) and clathrin (blue) (images are representatives of at least five captures). Note that there was no leading edge in the presence of this dynamin mutant, and endophilin-positive assemblies were distributed across the plasma membrane. Dynamin1-T65A

formed short tubules from the plasma membrane and at the tips of some of these tubules (arrowheads) in the boxed area (edges of cells) there were endophilin but not clathrin signals. **d**, Co-immunoprecipitation of  $\Delta$ H0-BAR-EGFP with endogenous endophilin (red box, left) or with Myc-tagged overexpressed endophilin (red box, right). **e**, Flow cytometry profiles of internalized transferrin (Alexa546-labelled) in cells overexpressing increasing amounts of EGFP-AP180-DN (red) or  $\Delta$ H0-BAR-EGFP (black).  $n = 2,000$  cells for each condition. **f**, Confocal images of cells overexpressing the indicated mutants (blue) for 24 h or treated with the indicated inhibitors for 5 min before stimulation with 10  $\mu$ M denopamine (4 min), fixed and immunostained for endogenous endophilin-A (green) and actin (phalloidin, red) (images are representative of at least six captures). Arrowheads point to EPAs. **g**, Live-cell confocal images before and after (+30 s) the addition of indicated inhibitors to cells expressing endophilin A2-RFP (green) and clathrin (EGFP-LCa, red) or dynamin (dynamin 2-EGFP, red) (images are representatives of at least three captures from independent experiments). Arrowheads point to endophilin spots. Scale bars: 10  $\mu$ m (**a**, **c**), 5  $\mu$ m (**f**, **g**) and 1  $\mu$ m (**c**, insets).

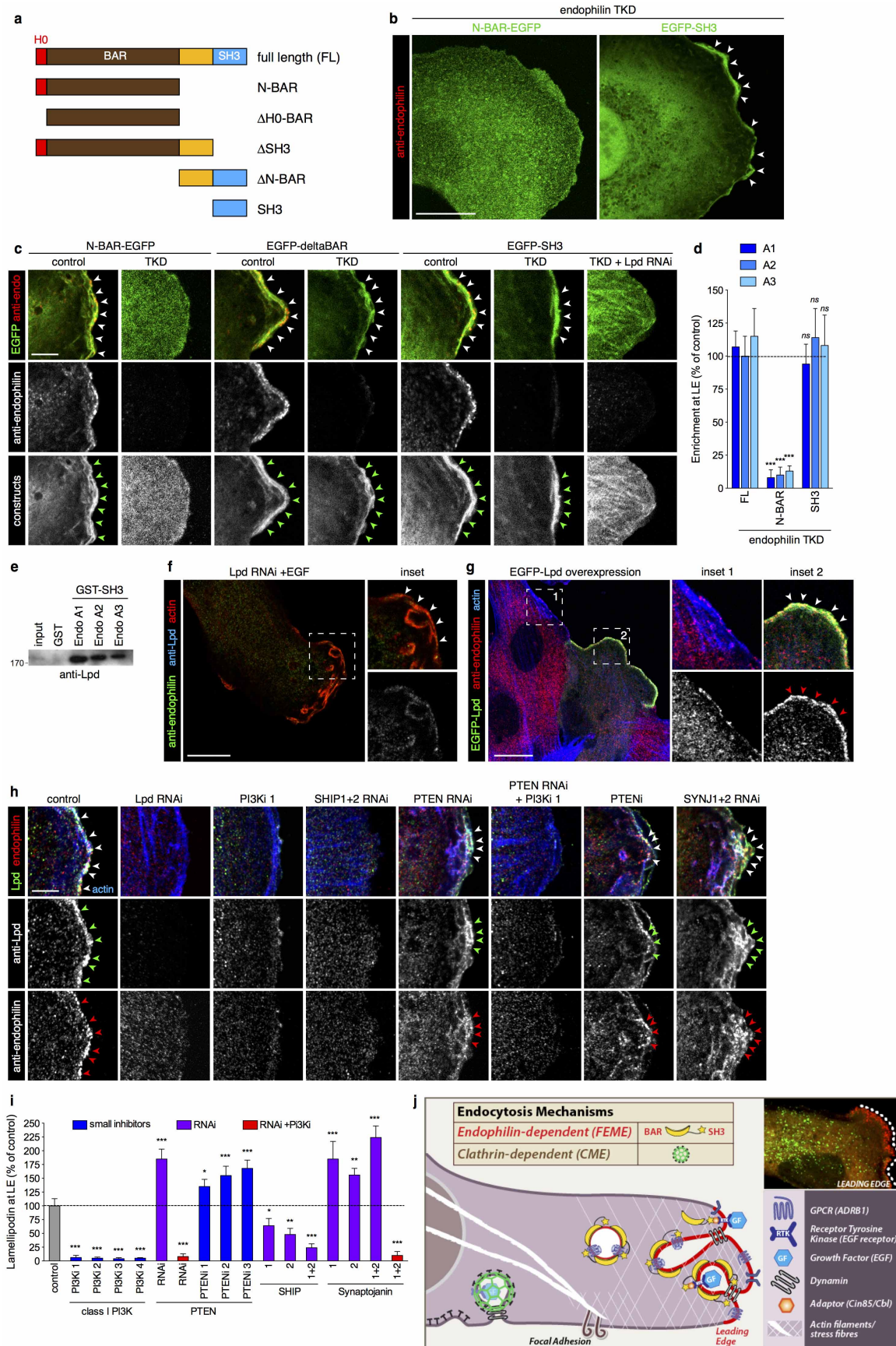




**Extended Data Figure 9 | The recruitment of endophilin at leading edges of cells requires PtdIns(3,4)P<sub>2</sub>.** Related to Fig. 5. **a**, Receptor uptake assay (HA-receptor-EGFP internalized, corrected for total levels) in cells treated with the depicted inhibitors (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **b**, Wound-closure assay. BSC1 and RPE1 cell monolayers were wounded and cell migration into the wound sites were assessed after 16 h. The inhibitors were added just after wounding (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). **c**, Confocal images of cells treated with the indicated inhibitors or RNAi, fixed and immunostained for endogenous endophilin-A (green) and actin (phalloidin, red) (images are representative of at least six captures from three independent experiments). Arrowheads point to endophilin foci at the leading edges of cells. **d**, Confocal images of cells overexpressing the indicated constructs fixed and immunostained for endogenous endophilin-A

(red) and actin (phalloidin, blue) (images are representative of at least six captures from three independent experiments). **e**, Quantifications of PtdIns(4,5)P<sub>2</sub> (mocha), PtdIns(3,4)P<sub>2</sub> (red) or PtdIns(3,4,5)P<sub>3</sub> (asparagus) levels from cells treated with the indicated siRNA and PI(3)K inhibitor, immunostained for the lipids and measured using a plate reader (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant; \* $P < 0.05$ ; \*\*\* $P < 0.001$ , one-way ANOVA and Dunnett's test versus respective control levels). **f**, Confocal images of cells treated with INP4A and INP4B or control siRNA, fixed and immunostained for endogenous endophilin-A (green) and LAMP-1 (red) (images are representative of at least six captures from three independent experiments). Intensity profiles were acquired along the indicated lines. Scale bars: 20  $\mu$ m (**d**), 10  $\mu$ m (**f**) and 5  $\mu$ m (**c**).





**Extended Data Figure 10 | Lamellipodin recruits endophilin at the leading edge.** Related to Fig. 5. **a**, Scheme depicting the endophilin constructs used. **b, c**, Confocal images of cells depleted for endogenous endophilin (TKD) or not (control siRNA) and expressing the indicated constructs, fixed and immunostained for endogenous endophilin to ascertain the depletion in the cells imaged (images are representatives of at least six captures from three independent experiments). **d**, Quantification of construct levels at the leading edges of cells depleted of endogenous endophilin (TKD) and expressing the indicated endophilin constructs (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant; \*\*\* $P < 0.001$  one-way ANOVA and Dunnett's test versus respective full-length levels). **e**, Pull-down experiments using GST-SH3 domains of endophilin-A1, 2 or 3 (or GST as a control) with EGFP-lamellipodin. 10% of cell extract was used as 'input'. **f**, Confocal image of cell depleted for lamellipodin (lpd) and treated with  $200 \text{ ng ml}^{-1}$  EGF for 5 min to stimulate PI(3)K and induce ruffles, fixed and immunostained for endophilin (green), actin (red) and lamellipodin (blue, to ascertain the depletion) (images are representatives of at least six captures from three independent experiments). **g**, Confocal image of a cell overexpressing

EGFP-lamellipodin next to a non-transfected cell, fixed and immunostained for endophilin (red) and actin (blue) (images representatives of at least six captures from three independent experiments). **h**, Confocal images of cells treated with the indicated RNAi and inhibitors, fixed and immunostained for lamellipodin (green), endophilin (red) and actin (blue) (images representatives of at least six captures from three independent experiments). **i**, Quantification of levels of endogenous lamellipodin at the leading edges of cells treated with the indicated small inhibitors (blue), RNAi (purple), RNAi plus class I PI(3)K inhibitor ('PI3Ki', GDC-0941, 50 nM for 5 min; red) (mean  $\pm$  s.e.m.,  $n = 3$  independent experiments; NS, non-significant; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  one-way ANOVA and Dunnett's test versus control). **j**, Model: endophilin-dependent (FEME) endocytosis from the leading edge. Endophilin-coated vesicles/tubules associate with receptors (after ligand binding) or receptor-adaptors (like CIN85/Cbl) via their SH3 domains, and promote membrane curvature with their N-terminal N-BAR domains. Vesicle scission is promoted by dynamin, which is also recruited by endophilin. Scale bars: 20  $\mu\text{m}$  (**b, g**), 10  $\mu\text{m}$  (**f**) and 5  $\mu\text{m}$  (**c, h**).

# Meikin is a conserved regulator of meiosis-I-specific kinetochore function

Jihye Kim<sup>1\*</sup>, Kei-ichiro Ishiguro<sup>1\*</sup>, Aya Nambu<sup>1</sup>, Bungo Akiyoshi<sup>1</sup>, Shihori Yokobayashi<sup>1</sup>, Ayano Kagami<sup>1</sup>, Tadashi Ishiguro<sup>1</sup>, Alberto M. Pendas<sup>2</sup>, Naoki Takeda<sup>3</sup>, Yogo Sakakibara<sup>4</sup>, Tomoya S. Kitajima<sup>4</sup>, Yuji Tanno<sup>1</sup>, Takeshi Sakuno<sup>1</sup> & Yoshinori Watanabe<sup>1</sup>

**The kinetochore is the crucial apparatus regulating chromosome segregation in mitosis and meiosis. Particularly in meiosis I, unlike in mitosis, sister kinetochores are captured by microtubules emanating from the same spindle pole (mono-orientation) and centromeric cohesion mediated by cohesin is protected in the following anaphase. Although meiotic kinetochore factors have been identified only in budding and fission yeasts, these molecules and their functions are thought to have diverged earlier. Therefore, a conserved mechanism for meiotic kinetochore regulation remains elusive. Here we have identified in mouse a meiosis-specific kinetochore factor that we termed MEIKIN, which functions in meiosis I but not in meiosis II or mitosis. MEIKIN plays a crucial role in both mono-orientation and centromeric cohesion protection, partly by stabilizing the localization of the cohesin protector shugoshin. These functions are mediated mainly by the activity of Polo-like kinase PLK1, which is enriched to kinetochores in a MEIKIN-dependent manner. Our integrative analysis indicates that the long-awaited key regulator of meiotic kinetochore function is Meikin, which is conserved from yeasts to humans.**

In mitosis, the establishment of sister chromatid cohesion is dependent on cohesin in S phase and maintained until metaphase when the sister chromatids are captured by spindle microtubules from opposite poles and aligned on the spindle equator. For the onset of anaphase, the anaphase-promoting complex (APC) triggers the degradation of securin, an inhibitory chaperone for separase that cleaves the cohesin subunit RAD21 and removes cohesin complex along the entire chromosome. This removal of cohesin triggers the separation of sister chromatids and their movement to opposite poles, a process called equational division<sup>1–3</sup>. However, during meiosis, the meiotic cohesin REC8 mainly replaces RAD21 along the entire chromosomes; one round of DNA replication is followed by two rounds of nuclear division, which results in four haploid nuclei or gametes (Fig. 1a).

In the first division of meiosis (meiosis I), homologous chromosomes connected by chiasmata are captured from the opposite poles, whereas sisters are captured from the same pole (mono-orientation). At the onset of anaphase I, REC8 cohesin is cleaved by separase along the arm regions, but protected at centromeres until metaphase II (refs 4–6). Thus, mono-orientation and centromeric cohesion protection are two hallmarks of meiotic kinetochore function, which are widely conserved among eukaryotic organisms<sup>7–9</sup> (Fig. 1a). There is increasing evidence that cohesion protection is mediated by the centromeric protein shugoshin (SGO) and its partner protein phosphatase 2A (PP2A)<sup>10–15</sup>, which antagonizes REC8 phosphorylation, a prerequisite of cleavage<sup>16,17</sup>. So far, meiosis-specific kinetochore proteins have been identified only in two yeasts (*Saccharomyces cerevisiae* Spo13 and Mam1 (monopolin subunit), and *Schizosaccharomyces pombe* Moa1)<sup>18–23</sup>. Puzzlingly, however, because their structural and functional similarities remain to be identified, conservation of meiotic kinetochore regulation is questionable even between yeasts<sup>8,9</sup>. Therefore, in this study, we address the long-standing question of whether meiotic kinetochore regulation is conserved from yeasts to mammals, and, if so, how.

## Mammalian meiotic kinetochore protein MEIKIN

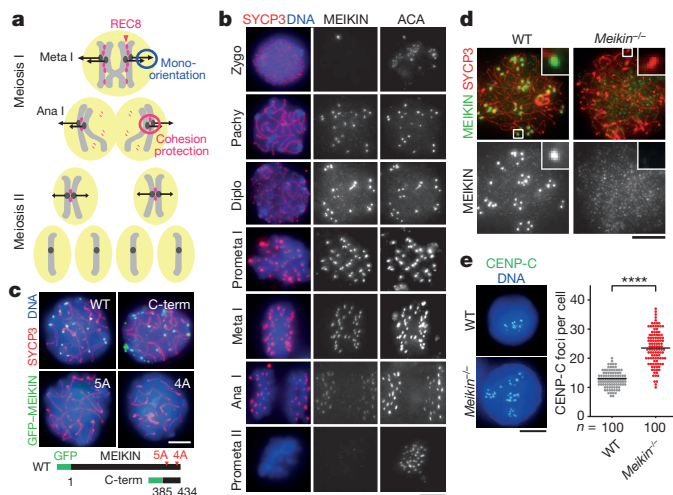
Fission yeast protein Moa1 interacts directly with the conserved kinetochore protein Cnp3 (CENP-C homologue), and localizes to the kinetochore in meiosis I (ref. 24). To identify an equivalent meiosis-specific kinetochore protein in mammals, we searched for proteins that interact with mouse CENP-C in a yeast two-hybrid assay using a cDNA library prepared from mouse testis (Extended Data Fig. 1a). The most frequently obtained clones (*4930404A10Rik* gene) encoded a novel protein, which we named MEIKIN (for meiosis-specific kinetochore protein). *Meikin* shows specific expression in germ cells (both testis and ovary) but not in other organs (Extended Data Fig. 1b). Immunoprecipitation assays using testis chromatin extracts indicate that MEIKIN indeed forms a complex with CENP-C (Extended Data Fig. 1c). Blast search analysis revealed that MEIKIN is a novel uncharacterized protein conserved among vertebrates (Extended Data Fig. 2).

To determine the localization of MEIKIN, we immunostained for MEIKIN in spermatocytes along with SYCP3, a component of the axial elements, and ACA (anti-centromeric antibodies), which stains constitutive centromeric proteins including CENP-C (Fig. 1b and Extended Data Fig. 3a). MEIKIN appears at centromeres during the pachytene stage when homologous chromosomes (homologues) are synapsed. Centromeric MEIKIN signals reach a peak during diplotene stage, persist with gradual reduction until metaphase I, and finally disappear in anaphase I. In meiosis II, MEIKIN does not reappear on chromatin. This localization contrasts with that of ACA (or CENP-C), which increase during zygotene and persist throughout meiosis I and meiosis II (Fig. 1b and Extended Data Fig. 3a). A similar localization pattern of MEIKIN was observed in oocytes (Extended Data Fig. 3b, c). To determine the mechanism of MEIKIN localization, we narrowed down the kinetochore localization sequences of MEIKIN by expressing GFP-fusion versions in testis, and identified carboxy terminus conserved sequences that have an essential role in localization to kinetochores (Fig. 1c and

<sup>1</sup>Laboratory of Chromosome Dynamics, Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Tokyo 113-0032, Japan. <sup>2</sup>Instituto de Biología Molecular y Celular del Cáncer (CSIC-USAL), 37007 Salamanca, Spain. <sup>3</sup>Center for Animal Resources and Development, Kumamoto University, 2-2-1 Honjo, Kumamoto 860-0811 Japan. <sup>4</sup>Laboratory for Chromosome Segregation, RIKEN Center for Developmental Biology, 2-2-3 Minatogima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan.

\*These authors contributed equally to this work.





**Figure 1 | Meiotic kinetochore protein MEIKIN.** **a**, Schematic drawing of behaviour of homologous chromosomes during meiosis, showing metaphase I (Meta I) and anaphase I (Ana I). **b**, Spermatocytes stained for MEIKIN, ACA, SYCP3 and DAPI (4',6-diamidino-2-phenylindole, DNA stain); zygotene (Zygo), pachytene (Pachy), diplotene (Diplo), prometaphase I (Prometa I) and prometaphase II (Prometa II). **c**, Spermatocytes from the wild-type testis transfected with GFP-tagged MEIKIN stained for GFP, SYCP3 and DAPI (DNA). Alanine substitutions (5A and 4A) were introduced on the C-terminal conserved sequences of MEIKIN (also see Extended Data Fig. 2). WT, wild type. **d**, Wild-type and *Meikin*<sup>-/-</sup> diplotene spermatocytes stained for MEIKIN and SYCP3. Inset, enlarged image of a kinetochore. **e**, Wild-type and *Meikin*<sup>-/-</sup> round spermatids stained for CENP-C and DAPI (DNA). \*\*\*\**P* < 0.0001, unpaired *t*-test. Scale bars, 5  $\mu$ m.

Extended Data Fig. 1d and 2). We conclude that MEIKIN is a meiosis-I-specific kinetochore protein whose localization is mediated by its interaction with CENP-C.

### *Meikin*<sup>-/-</sup> mice are infertile

To determine the function of MEIKIN in meiosis, we generated *Meikin*<sup>-/-</sup> mice (Extended Data Fig. 4a–c). Immunostaining of spermatocytes confirmed that MEIKIN proteins are absent from the kinetochores in *Meikin*<sup>-/-</sup> (Fig. 1d). Although homozygous *Meikin*<sup>-/-</sup> mice develop normally and exhibited no overt phenotype, both male and female KO (knockout) mice are completely infertile (data not shown). Cytological inspection of testis and ovary detected no obvious difference between wild-type and *Meikin*<sup>-/-</sup> mice, although mature sperms are rarely observed in the epididymis of *Meikin*<sup>-/-</sup> mice (Extended Data Fig. 4d–g). Immunostaining of spermatocytes revealed apparently normal meiotic prophase progression as the structures of the axial elements are properly formed and proceeded into synapsis in pachytene stage and further into condensed meiotic chromosomes (Fig. 1d, see below). The round spermatids produced after meiosis, however, become enlarged, because meiosis II did not appear to have occurred (Fig. 1e). These results suggest that infertility in male *Meikin*<sup>-/-</sup> mice originates from defects in meiotic chromosome segregation and subsequent spermatogenesis rather from a problem with meiotic prophase (see below).

### MEIKIN protects centromeric cohesion

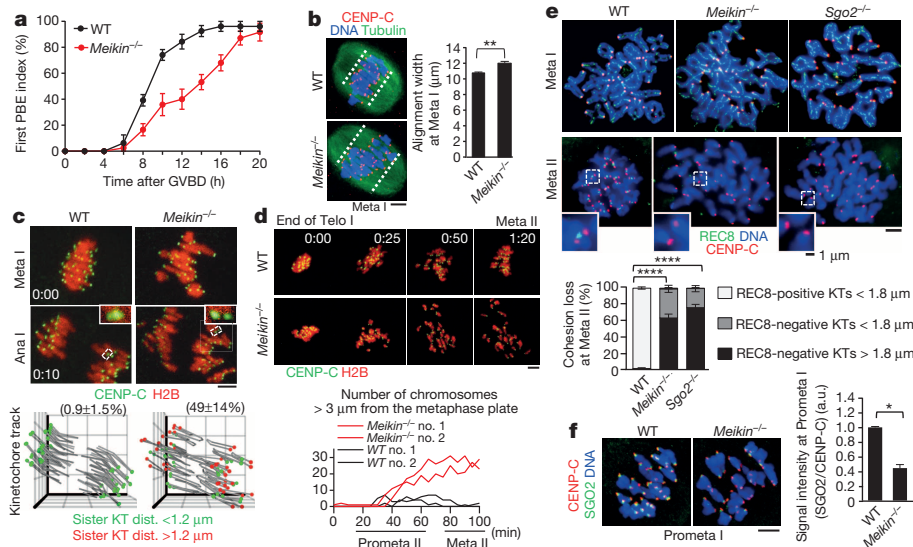
Similar to wild-type mice, female *Meikin*<sup>-/-</sup> mice produced germinal vesicle (GV) stage oocytes presumably as a result of normal prophase progression. We examined meiotic division by isolating GV-stage oocytes. When placed in culture medium, the GV-stage oocytes resume meiosis, undergo germinal vesicle breakdown (GVBD), form meiosis I spindles and align bivalent chromosomes using their chiasmata. When all bivalents are aligned and come under tension, chiasmata are resolved and homologues are segregated to opposite poles. Because one set of chromosomes is retained in the oocytes but the other is segregated into the

first polar body, the meiosis I division can be monitored by the extrusion of the first polar body. Light microscopy observation revealed that *Meikin*<sup>-/-</sup> oocytes extrude the first polar body on average 2 h later than oocytes from wild-type littermates, although most mutant oocytes (>90%) eventually enter meiosis I division (Fig. 2a). In wild-type oocytes, bivalents are aligned by 6 h after GVBD at the spindle equator, thus forming a metaphase plate (Fig. 2b). Although *Meikin*<sup>-/-</sup> oocytes also show bivalent alignment in metaphase I, chromosome congression is partly perturbed (Fig. 2b). Because inactivation of the spindle assembly checkpoint cancels the delay in anaphase onset of *Meikin*<sup>-/-</sup> oocytes compared to wild-type oocytes (Extended Data Fig. 5), we reasoned that chromosome alignment defects and consequent activation of the spindle assembly checkpoint might be the reason for the delay in the onset of anaphase I.

To analyse the spatiotemporal dynamics of kinetochores and chromosomes during meiosis I, we recorded four-dimensional (4D) data sets of kinetochores and chromosomes labelled with 2mEGFP–CENP-C and histone 2B (H2B)–mCherry<sup>25</sup>. We followed the frames at the onset of anaphase in wild-type and *Meikin*<sup>-/-</sup> oocytes and confirmed that the bivalents aligned in the metaphase plate are indeed separated in anaphase I in both oocytes (Fig. 2c). Strikingly, however, many pairs of sister kinetochores start to split at the onset of anaphase in *Meikin*<sup>-/-</sup> oocytes but rarely in wild-type oocytes (Fig. 2c). Accordingly, chromosome alignment in metaphase II is largely disrupted in *Meikin*<sup>-/-</sup> oocytes (Fig. 2d), presumably because sister chromatids were prematurely separated. Indeed, immunostaining of spread chromosomes revealed that although REC8 localizes normally to the interchromatid axes on metaphase I bivalents, centromeric REC8 signals are lost from metaphase II chromosomes in *Meikin*<sup>-/-</sup> oocytes (Fig. 2e). The loss of centromeric cohesin REC8 (and cohesion) during anaphase I is reminiscent of the phenotype of mice defective in shugoshin-2 (SGO2, also known as SGOL2), the protector of centromeric cohesin in mouse meiosis<sup>11,12</sup> (Fig. 2e). We then examined bivalents for SGO2 localization, which usually appears at centromeres in metaphase I (refs 11, 12). Intriguingly, SGO2 localization is diminished in *Meikin*<sup>-/-</sup> oocytes (Fig. 2f), possibly explaining the impaired protection of centromeric cohesion in anaphase I. Similar defects were observed in *Meikin*<sup>-/-</sup> spermatocytes (Extended Data Fig. 6a). Notably, kinetochore separation defects during anaphase I are less severe in *Meikin*<sup>-/-</sup> spermatocytes as compared with *Sgo2*<sup>-/-</sup> spermatocytes (Extended Data Fig. 6b), indicating that the protection defects caused by *Meikin*<sup>-/-</sup> are somewhat weaker than those caused by *Sgo2*<sup>-/-</sup>.

### MEIKIN facilitates mono-orientation

Because meiosis I chromosome alignment is partly perturbed in *Meikin*<sup>-/-</sup> oocytes (Fig. 2b), we postulated that mono-orientation of sister kinetochores might be also impaired in the mutant mice, similar to the way it is impaired in fission yeast *moa1Δ* cells<sup>23</sup>. We consistently found that the sister kinetochore pairs distance of prometaphase I chromosomes is significantly increased (~20%) in *Meikin*<sup>-/-</sup> mice as compared to wild-type mice (Fig. 3a and Extended Data Fig. 6c). Given that mono-orientation defects are significantly suppressed by chiasmata (or tension between homologues)<sup>23,26</sup>, we sought to use *Mlh1*<sup>-/-</sup> mice that have severely reduced chiasmata and accumulate univalent chromosomes in meiosis I. Because of the intrinsic property of mono-orientation of sister kinetochores, most univalents in *Mlh1*<sup>-/-</sup> oocytes fail to align on the metaphase plate and mainly inhibit the onset of anaphase I (ref. 27) (Fig. 3b). Interestingly, further depletion of MEIKIN in an *Mlh1*<sup>-/-</sup> background alleviates the anaphase I inhibition, although anaphase I onset is delayed in *Meikin*<sup>-/-</sup> when chiasmata are present (Figs 2a and 3b). We examined chromosome alignment at 10 h after GVBD when most *Mlh1*<sup>-/-</sup> or *Mlh1*<sup>-/-</sup>*Meikin*<sup>-/-</sup> oocytes have not yet entered anaphase I. Univalents congressed at the metaphase plate and sister kinetochore pairs separated by >0.6  $\mu$ m were seen more frequently in *Mlh1*<sup>-/-</sup>*Meikin*<sup>-/-</sup> than *Mlh1*<sup>-/-</sup> (Fig. 3c). Thus we infer that mono-orientation is impaired in *Mlh1*<sup>-/-</sup> *Meikin*<sup>-/-</sup> oocytes. In contrast, *Mlh1*<sup>-/-</sup>



**Figure 2 | MEIKIN is required for the protection of centromeric cohesion during meiosis I.** **a**, Cumulative oocyte first polar body extrusion (PBE) rates after GV breakdown (GVBD). Mean  $\pm$  s.e.m. from 4 independent experiments, total oocytes were  $n = 40$  for wild-type,  $n = 58$  for *Meikin*<sup>-/-</sup>. **b**, Metaphase I oocytes (6 h after GVBD) stained for  $\alpha$ -tubulin, CENP-C and DAPI (DNA), and the chromosome alignment widths were measured (mean  $\pm$  s.e.m. of three independent experiments). A total of 10 oocytes were used in each experiment. **c**, Time-lapse imaging of the first meiotic division (h:mm) in oocytes expressing 2mEGFP-CENP-C and H2B-mCherry. The z-projection images at metaphase I and anaphase I (top). The images were reconstructed in 3D, and kinetochore tracks are indicated by grey lines (bottom). The sister kinetochore distances at anaphase I are colour-coded as indicated. The ratios of separated sister kinetochores (distance  $>1.2 \mu\text{m}$ ) at anaphase I were measured in 3 oocytes (mean  $\pm$  s.d.). Also see Supplementary Video 1. **d**, Time-lapse imaging of the second meiotic division (h:mm) in

oocytes as in **c** (top). Resumption of kinetochore movement after first PBE was assumed as the end of telophase I (Telo I). The number of kinetochores with distance  $>3 \mu\text{m}$  from the metaphase plate were counted over time (bottom). Also see Supplementary Video 2. **e**, Oocytes at metaphase I and metaphase II stained for REC8, CENP-C and DAPI (DNA). The spread chromosomes at metaphase II were classified according to REC8 signals and kinetochore distance (mean  $\pm$  s.e.m. of three independent experiments). A total of 5 oocytes were used in each experiment. **f**, Prometaphase I oocytes (4 h after GVBD) stained for SGO2, CENP-C and DAPI (DNA) in whole mount. The signal intensity of SGO2 adjacent to the centromere was quantified and normalized to that of CENP-C (mean  $\pm$  s.e.m. of three independent experiments). In each experiment, 15 centromeres from an oocyte were quantified ( $n = 5$  cells). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ , unpaired *t*-test (**b**, **e**, **f**). Scale bars,  $5 \mu\text{m}$ .

*Sgo2*<sup>-/-</sup> oocytes did not show such defects (Fig. 3c). A similar mono-orientation defect was observed in *Mlh1*<sup>-/-</sup> *Meikin*<sup>-/-</sup> spermatocytes (Fig. 3d). Therefore, we conclude that MEIKIN plays a crucial role in mono-orientation in addition to its role in protecting centromeric cohesion.

## MEIKIN recruits PLK1

To elucidate the molecular function of MEIKIN, we searched for MEIKIN-interacting proteins in a yeast two-hybrid system, as well as by immunoprecipitation with mass spectrometry. In both assays, we obtained Polo-like kinase PLK1 in addition to CENP-C (Extended Data Fig. 7). Immunoprecipitation assays indicated that MEIKIN co-precipitates PLK1 in testis chromatin extracts (Fig. 4a). Immunostaining using spermatocytes indicated that PLK1 localizes at kinetochores in diplotene, peaks in metaphase I and declines during anaphase I, when MEIKIN disappears completely. In metaphase II, centrosome localization is prominent but kinetochore localization is markedly decreased to levels comparable to those in mitosis, indicating that the kinetochore enrichment of PLK1 is specific to meiosis I (Fig. 4b). Strikingly, PLK1 signals at kinetochores were diminished in *Meikin*<sup>-/-</sup> oocytes and mainly dispersed in *Meikin*<sup>-/-</sup> spermatocytes (Fig. 4b, c). These results indicate that MEIKIN plays a crucial role in enriching PLK1 to kinetochores in diplotene and in sustaining its localization at least until anaphase I.

## PLK1 acts in protection and mono-orientation

These results raised the possibility that PLK1 activity enriched to kinetochores might be required for kinetochore regulation in meiosis I. To explore this further, we sought to deplete PLK1 activity during meiosis I by adding the PLK1 inhibitor BI 2536 (ref. 28) to the oocyte culture.

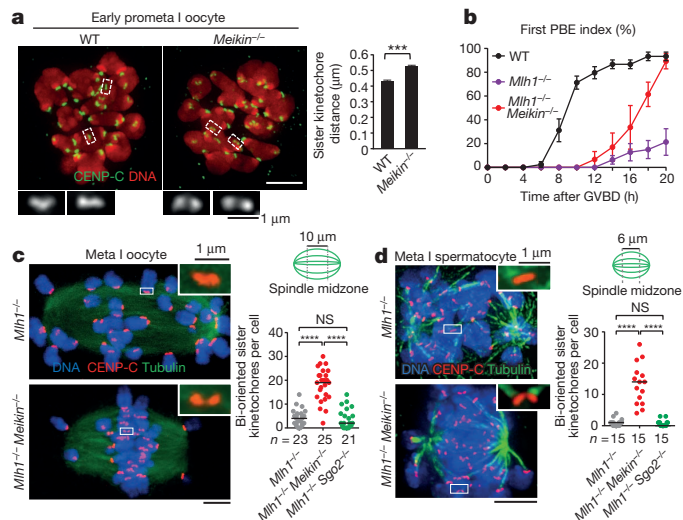
We treated oocytes with BI 2536 for 2 h between 4 h and 6 h after GVBD, because this transient treatment allowed  $\sim 70\%$  of oocytes to undergo meiosis I division and suppressed PLK1 activity at kinetochores (Extended Data Fig. 8a, b). More than half of BI 2536 treated oocytes showed chromosome misalignment in metaphase II, most likely because partial or complete sister chromatid separation happen during anaphase I (Fig. 4d and Extended Data Fig. 8c). Accordingly, centromeric REC8 on metaphase II chromosomes was mainly diminished in the BI 2536 treated oocytes (Fig. 4e and Extended Data Fig. 8d). These results suggest that PLK1 activity is required, at least after prometaphase I, for the protection of centromeric REC8 cohesin in the following anaphase I.

We next examined whether PLK1 activity is also required for the mono-orientation of sister kinetochores in meiosis I by treating *Mlh1*<sup>-/-</sup> oocytes with BI 2536 between 9 h and 10 h after GVBD. Although univalents largely failed to congress in *Mlh1*<sup>-/-</sup> oocytes, BI 2536 treatment significantly increased the numbers of univalents congressed at the metaphase plate and sister kinetochore pairs separated by  $>0.6 \mu\text{m}$  (Fig. 4f and Extended Data Fig. 8e). Thus, we conclude that PLK1 also plays a role in promoting mono-orientation in oocytes.

## Conservation of MEIKIN

Our analyses suggested that mouse MEIKIN is analogous to *S. pombe* Moa1 because both mutants show defects in mono-orientation and the protection of centromeric cohesion, albeit to different extents<sup>23</sup>. Accordingly, we have found that, like MEIKIN, Moa1 carries a putative polo-box domain (PBD)-binding site (Ser-Thr-Pro)<sup>29</sup> (Fig. 5a), although overall amino acid sequences are not conserved. Immunoprecipitation assays revealed that Moa1 associates with the fission yeast Polo-like kinase Plol through the PBD-binding motif *in vivo* (Fig. 5b). Because Plol is



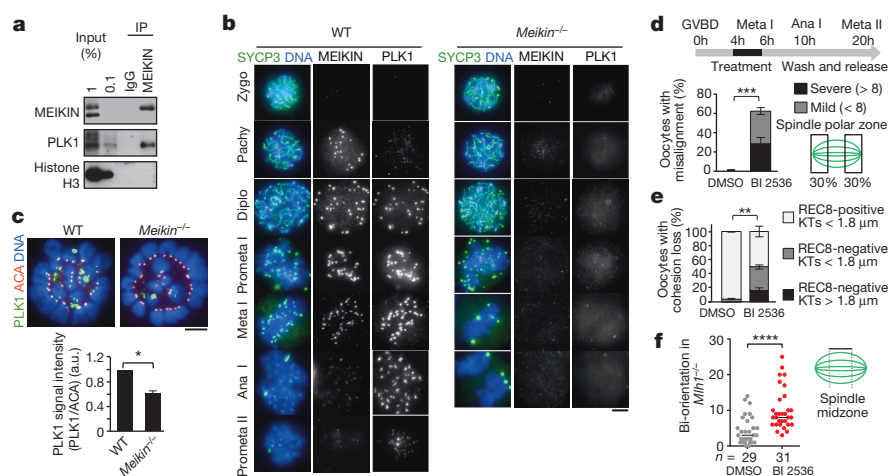


**Figure 3 | MEIKIN regulates mono-orientation.** **a**, Wild-type and *Meikin*<sup>-/-</sup> oocytes at early prometaphase I (2 h after GVBD) stained for CENP-C and DAPI (DNA). The pairs of sister kinetochores are magnified. Measurement of the distance between two CENP-C signals of a sister kinetochore pair on different z planes (mean + s.e.m. of three independent experiments). In each experiment, 10 kinetochores were measured in a cell ( $n = 5$  cells). **b**, Cumulative first PBE rates after GVBD (mean ± s.e.m. of three independent experiments). Total oocytes were  $n = 45$  for wild type,  $n = 29$  for *Mlh1*<sup>-/-</sup> and  $n = 29$  for *Mlh1*<sup>-/-</sup> *Meikin*<sup>-/-</sup>. **c**, Oocytes at 10 h after GVBD stained for  $\alpha$ -tubulin, CENP-C and DAPI (DNA) in whole mount. Bi-oriented sister kinetochores; distance >0.6  $\mu$ m with horizontal angle (0–10°). The number of bi-oriented sister kinetochores at the spindle midzone (10  $\mu$ m width centred between spindle poles; only spindles up to 30  $\mu$ m in length were used) was scored in each cell. Note that the spindle length becomes longer in *Mlh1*<sup>-/-</sup> because of less alignment of chromosomes<sup>50</sup>. **d**, Metaphase I spermatocytes were stained and examined as in **c**. Spindle midzone, 6  $\mu$ m (only spindles up to 15  $\mu$ m in length were used). Sister kinetochores distance >0.6  $\mu$ m as 'bi-oriented'. NS, not significantly different. \*\*\*  $P < 0.001$ , \*\*\*\*  $P < 0.0001$ , unpaired  $t$ -test (**a**, **c**, **d**). Scale bars, 5  $\mu$ m (unless otherwise indicated). *Mlh1*<sup>-/-</sup> *Meikin*<sup>-/-</sup> and *Mlh1*<sup>-/-</sup> *Sgo2*<sup>-/-</sup> were compared to *Mlh1*<sup>-/-</sup> with C57BL/6 background littermates (**c**, **d**).

indeed localized to kinetochores in meiosis I (ref. 30) (Fig. 5c), we examined the functional relationship between Plo1 and Moa1. Meiosis-I-specific enrichment of Plo1 to kinetochores was abolished in cells carrying a mutation in the PBD-binding motif of Moa1 (*moa1-101A*) as well as in *moa1* $\Delta$  cells (Fig. 5c). We then examined the effects of these mutations on meiotic chromosome segregation in recombination-deficient (*rec12* $\Delta$ ) zygotes, in which chiasmata are not formed. The results indicate that *moa1-101A* cells undergo equational segregation of sister chromatids in meiosis I, as in *moa1* $\Delta$  cells (Fig. 5d), stressing the importance of the PBD-binding motif for Moa1 function.

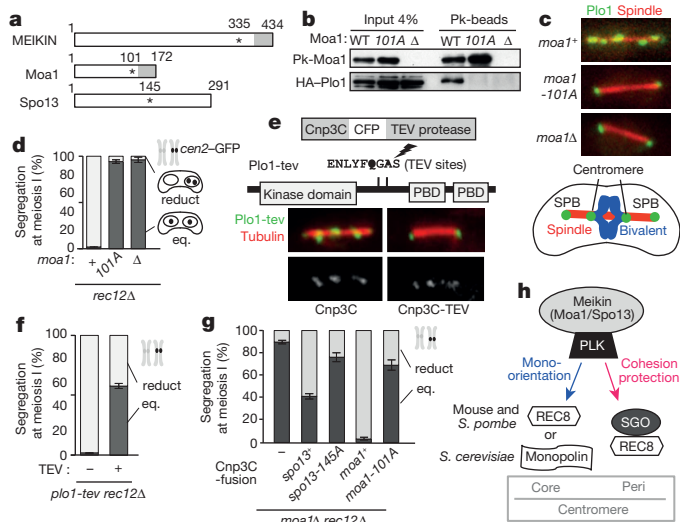
To delineate the role of Plo1 upon Moa1, we sought to inactivate Plo1 in meiosis. Because a reduction in Plo1 at the cellular level impairs spindle formation and nuclear division (data not shown), we sought to mutate endogenous Plo1 to make it cleavable by tobacco etch virus (TEV) protease (Plo1-tev) and thus inactivate it only around the kinetochore by co-expressing Cnp3C-TEV (TEV protease fused with a kinetochore-localizing peptide)<sup>23</sup>. Indeed, *plo1-tev* cells expressing Cnp3C-TEV showed a reduction of Plo1-tev signals selectively at kinetochores but not at SPBs (yeast centrosomes) in metaphase I, whereas *plo1-tev* cells expressing a control Cnp3C not fused to TEV showed normal localization of Plo1-tev (Fig. 5e). Strikingly, *plo1-tev* cells expressing Cnp3C-TEV mainly underwent equational segregation at meiosis I (Fig. 5f). Thus, we conclude that Plo1 recruited by Moa1 to kinetochores fulfils a similar crucial role in promoting mono-orientation and centromeric cohesion protection in fission yeast as it does in mice.

These findings are reminiscent of the observation that budding yeast Spo13 possesses the PBD-binding motif and cooperates with PLK (Cdc5)<sup>31</sup>, although overall amino acid sequence homology between Moa1 and Spo13 is missing, and a meiotic role for Cdc5 in the protection of cohesin is a subject of debate<sup>16,32,33</sup> (Fig. 5a). To examine the functional conservation between Spo13 and Moa1, we expressed Spo13 fused with the kinetochore targeting peptide Cnp3C in *moa1* $\Delta$  cells (Spo13 itself cannot localize to fission yeast kinetochores; data not shown). Strikingly, the mono-orientation defect of *moa1* $\Delta$  was significantly suppressed by Spo13-Cnp3C, but not by Spo13-145A-Cnp3C, which has a mutation in the PBD-binding motif<sup>31</sup> (Fig. 5g). Similar results were obtained by expressing Moa1 and its mutant version fused to Cnp3C in *moa1* $\Delta$  cells, although Moa1-Cnp3C produced a more



**Figure 4 | PLK1 is required for mono-orientation and the protection of centromeric cohesion.** **a**, Immunoprecipitates from mouse testis chromatin extracts. **b**, Spermatocytes stained for MEIKIN, PLK1, SYCP3 and DAPI (DNA). **c**, Prometaphase I oocytes (3 h after GVBD) stained for PLK1, ACA and DAPI (DNA). The relative intensities of PLK1 normalized to that of ACA (mean + s.e.m. of three independent experiments). In each experiment, 15 kinetochores from an oocyte were quantified ( $n = 4$  cells). **d**, A schematic time course of PLK1 inhibition experiment in wild-type oocytes (top). The number of single chromatids at the spindle polar zone was scored at metaphase II (bottom). Mean ± s.e.m. of three independent experiments. Ten oocytes were counted in each experiment. **e**, The spread chromosomes at metaphase II were classified as in Fig. 2e. Mean ± s.e.m. of three independent experiments. 5 oocytes were counted in each experiment.  $P$  values, unpaired  $t$ -test from REC8-negative categories. **f**, *Mlh1*<sup>-/-</sup> oocytes (C57BL/6 background) were treated with BI 2536 between 9–10 h after GVBD and bi-oriented sister kinetochores were scored as indicated in Fig. 3c. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , unpaired  $t$ -test (**c–f**). Scale bars, 5  $\mu$ m.





**Figure 5 | *S. pombe* Moa1 and *S. cerevisiae* Spo13 are functional homologues of MEIKIN.** **a**, Schematic representation of mouse MEIKIN, *S. pombe* Moa1 and *S. cerevisiae* Spo13. Asterisks indicate the polo-box binding motif (STP). Shaded boxes indicate the CENP-C binding region. **b**, Co-precipitation of Moa1 and Plo1 in *S. pombe* meiotic cell extracts prepared from the indicated cells expressing Pk-Moa1 and HA-Plo1. **c**, Plo1-GFP localizes at kinetochores and SPB in metaphase I, while the kinetochore localization is abolished in *moa1-101A* and *moa1Δ* cells. The spindles were visualized by expressing cyan fluorescent protein (CFP)-Atb2 ( $\alpha 2$ -tubulin). Schematic representation of a metaphase I zygote is shown. **d**, The indicated zygotes examined for reductional (reduct) or equational (eq.) segregation of heterozygous *cen2*-GFP during meiosis I. **e**, Schematic diagram of Cnp3C-TEV and Plo1-tev, in which two TEV protease-recognition sites were inserted between the kinase domain and polo-box domain (PBD). Representative images of GFP-Plo1-tev signals along the metaphase I spindle (mCherry-Atb2) in cells expressing Cnp3C-CFP or Cnp3C-CFP-TEV. **f**, Segregation of heterozygous *cen2*-GFP during meiosis I was examined in *p101-tev rec12Δ* zygotes with vector (–) or expressing Cnp3C-CFP-TEV (+). **g**, Segregation of heterozygous *cen2*-GFP during meiosis I was examined in the *moa1Δ rec12Δ* zygotes expressing the indicated Cnp3C-fusion proteins. **h**, Schematic depiction of meiotic kinetochore regulation by Meikin, which cooperates with PLK. Error bars, s.d. from 3 independent experiments (**d**, **f**, **g**). Fission yeast strains used in this work are listed in Supplementary Table 1.

robust suppression. These results strongly suggest that Spo13 is the functional homologue of Moa1.

## Discussion

Motivated by our previous finding that *S. pombe* meiotic kinetochore regulator Moa1 binds to CENP-C<sup>24</sup>, we have identified a meiosis-specific kinetochore protein MEIKIN in mice. Although there is no significant sequence homology between MEIKIN and Moa1<sup>23</sup>, our study reveals striking biochemical and functional similarities between these two factors. Furthermore, these proteins share some properties with budding yeast Spo13 (refs 21, 22, 26, 31, 34). Our functional assays in fission yeast demonstrated that centromere-tethered Spo13 suppresses the *moa1Δ* defects dependent on its binding to PLK. These results are consistent with and support the notion that budding yeast Cdc5 is required for mono-orientation and for protecting cohesin at centromeres during meiosis I (refs 16, 32). Taken together, our results identify MEIKIN, Moa1 and Spo13 as functional homologues that form a novel protein family, for which we propose the name 'Meikin' (Fig. 5h).

Studies in budding yeast identified monopolin, which acts downstream of Spo13 to assemble a single kinetochore from one pair of centromeres<sup>20–22,35–37</sup>, however, a fission yeast monopolin homologue is dispensable for mono-orientation<sup>38,39</sup>. Instead, the fission yeast cohesin Rec8 plays a central role in mono-orientation downstream of Moa1 by establishing cohesion at the core centromeres and thereby conjoining

sister kinetochores<sup>23,40</sup>, a mechanism that might be lacking in budding yeast<sup>20,35</sup> (Fig. 5h). In mouse oocytes, as in fission yeast, the artificial cleavage of REC8 cohesin at the core centromeres disrupts mono-orientation<sup>23,41</sup>, implying that REC8 is acting for mono-orientation downstream of MEIKIN and this might be conserved in plants as well<sup>9</sup> (Fig. 5h). As fission yeast Moa1 has been shown to be required only after DNA replication<sup>42</sup>, MEIKIN starts to localize at kinetochores in the pachytene stage, so MEIKIN may also regulate the maintenance of cohesion rather than its establishment. Consistent with these results, our analyses show that sister kinetochores on prometaphase I chromosomes are conjoined in wild-type mice but split in *Meikin*<sup>–/–</sup> mice, although they do not separate as much as mitotic prometaphase chromosomes (Fig. 3a and Extended Data Fig. 6c). Accordingly, the mono-orientation defects are weak so that they can be detected in univalents in *Meikin*<sup>–/–</sup> *Mlh1*<sup>–/–</sup> mice but are hard to detect in *Meikin*<sup>–/–</sup> bivalents (Figs 2b and 3c). We therefore assume that in addition to sister-kinetochore association promoted by MEIKIN, meiotic pericentric chromatin including REC8 cohesin may also structurally contribute to restrain sister kinetochore separation, and thereby to bias to mono-orientation.

We showed that MEIKIN promotes the protection of centromeric cohesion at least partly by stabilizing SGO2 at the centromeres, and that the defects in cohesion protection are somewhat weaker in *Meikin*<sup>–/–</sup> mice as compared with *Sgo2*<sup>–/–</sup> mice. Thus, the protection of centromeric cohesion is not entirely abolished in *Meikin*<sup>–/–</sup> mice, and intriguingly this is also the case in fission yeast *moa1Δ* and budding yeast *spo13Δ* cells<sup>21,23</sup>. These results suggest that Meikin might be a regulator of shugoshin-dependent protection pathway rather than a bona fide protector of cohesin (Fig. 5h). Although Meikin acts to enrich PLK to kinetochores, Meikin might not be a mere recruiter of PLK. In fact, PLK is localized to kinetochores in mitotic cells<sup>43,44</sup>, albeit less so than in meiotic cells, and PLK reduction is mild in *Meikin*<sup>–/–</sup> oocytes compared to spermatocytes, although meiotic kinetochore functions are similarly defective in females and males. Therefore, we assume that Meikin may play additional roles such as in PLK substrate recognition and have some sort of regulatory function. Our studies in mouse together with those in yeast provide a unified model, in which the conserved key regulator Meikin cooperates with PLK to promote mono-orientation and centromeric cohesion protection, two hallmarks of meiotic kinetochore function in eukaryotes. The current study thus provides a clear answer to the longstanding question of how these two distinct kinetochore functions, both essential for meiosis I, are coordinated at the molecular levels. Because MEIKIN is conserved in humans (Extended Data Fig. 9), MEIKIN, like REC8 cohesin<sup>45–48</sup>, might be a candidate influencing age-associated chromosome segregation errors, a major cause of human birth defects<sup>49</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 31 July; accepted 19 November 2014.

Published online 24 December 2014.

- Uhlmann, F., Lottspeich, F. & Nasmyth, K. Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* **400**, 37–42 (1999).
- Peters, J. M., Tedeschi, A. & Schmitz, J. The cohesin complex and its roles in chromosome biology. *Genes Dev.* **22**, 3089–3114 (2008).
- Nasmyth, K. & Haering, C. H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
- Buonomo, S. B. *et al.* Disjunction of homologous chromosomes in meiosis I depends on proteolytic cleavage of the meiotic cohesin Rec8 by separin. *Cell* **103**, 387–398 (2000).
- Kitajima, T. S., Miyazaki, Y., Yamamoto, M. & Watanabe, Y. Rec8 cleavage by separase is required for meiotic nuclear divisions in fission yeast. *EMBO J.* **22**, 5643–5653 (2003).
- Tachibana-Konwalski, K. *et al.* Rec8-containing cohesin maintains bivalents without turnover during the growing phase of mouse oocytes. *Genes Dev.* **24**, 2505–2516 (2010).
- Moore, D. P. & Orr-Weaver, T. L. Chromosome segregation during meiosis: building an unambivalent bivalent. *Curr. Top. Dev. Biol.* **37**, 263–299 (1997).

8. Brar, G. A. & Amon, A. Emerging roles for centromeres in meiosis I chromosome segregation. *Nature Rev. Genet.* **9**, 899–910 (2008).
9. Watanabe, Y. Geometry and force behind kinetochore orientation: lessons from meiosis. *Nature Rev. Mol. Cell Biol.* **13**, 370–382 (2012).
10. Kitajima, T. S., Kawashima, S. A. & Watanabe, Y. The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature* **427**, 510–517 (2004).
11. Lee, J. *et al.* Unified mode of centromeric protection by shugoshin in mammalian oocytes and somatic cells. *Nature Cell Biol.* **10**, 42–52 (2008).
12. Llano, E. *et al.* Shugoshin-2 is essential for the completion of meiosis but not for mitotic cell division in mice. *Genes Dev.* **22**, 2400–2413 (2008).
13. Kitajima, T. S. *et al.* Shugoshin collaborates with protein phosphatase 2A to protect cohesin. *Nature* **441**, 46–52 (2006).
14. Riedel, C. G. *et al.* Protein phosphatase 2A protects centromeric sister chromatid cohesion during meiosis I. *Nature* **441**, 53–61 (2006).
15. Marston, A. L., Tham, W. H., Shah, H. & Amon, A. A genome-wide screen identifies genes required for centromeric cohesion. *Science* **303**, 1367–1370 (2004).
16. Katis, V. L. *et al.* Rec8 phosphorylation by casein kinase 1 and Cdc7-Dbf4 kinase regulates cohesin cleavage by separase during meiosis. *Dev. Cell* **18**, 397–409 (2010).
17. Ishiguro, T., Tanaka, K., Sakuno, T. & Watanabe, Y. Shugoshin-PP2A counteracts casein-kinase-1-dependent cleavage of Rec8 by separase. *Nature Cell Biol.* **12**, 500–506 (2010).
18. Hugerat, Y. & Simchen, G. Mixed segregation and recombination of chromosomes and YACs during single-division meiosis in *spo13* strains of *Saccharomyces cerevisiae*. *Genetics* **135**, 297–308 (1993).
19. Klapholz, S. & Esposito, R. E. Recombination and chromosome segregation during the single division meiosis in *spo12-1* and *spo13-1* diploids. *Genetics* **96**, 589–611 (1980).
20. Tóth, A. *et al.* Functional genomics identifies monopolin: a kinetochore protein required for segregation of homologs during meiosis I. *Cell* **103**, 1155–1168 (2000).
21. Katis, V. L. *et al.* Spo13 facilitates monopolin recruitment to kinetochores and regulates maintenance of centromeric cohesion during yeast meiosis. *Curr. Biol.* **14**, 2183–2196 (2004).
22. Lee, B. H., Kiburz, B. M. & Amon, A. Spo13 maintains centromeric cohesion and kinetochore coorientation during meiosis I. *Curr. Biol.* **14**, 2168–2182 (2004).
23. Yokobayashi, S. & Watanabe, Y. The kinetochore protein Moa1 enables cohesion-mediated monopolar attachment at meiosis I. *Cell* **123**, 803–817 (2005).
24. Tanaka, K., Chang, H. L., Kagami, A. & Watanabe, Y. CENP-C functions as a scaffold for effectors with essential kinetochore functions in mitosis and meiosis. *Dev. Cell* **17**, 334–343 (2009).
25. Kitajima, T. S., Ohsugi, M. & Ellenberg, J. Complete kinetochore tracking reveals error-prone homologous chromosome biorientation in mammalian oocytes. *Cell* **146**, 568–581 (2011).
26. Shonn, M. A., McCarroll, R. & Murray, A. W. Spo13 protects meiotic cohesin at centromeres in meiosis I. *Genes Dev.* **16**, 1659–1671 (2002).
27. Woods, L. M. *et al.* Chromosomal influence on meiotic spindle assembly: abnormal meiosis I in female *Mlh1* mutant mice. *J. Cell Biol.* **145**, 1395–1406 (1999).
28. Lénárt, P. *et al.* The small-molecule inhibitor BI 2536 reveals novel insights into mitotic roles of polo-like kinase 1. *Curr. Biol.* **17**, 304–315 (2007).
29. Elia, A. E. *et al.* The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell* **115**, 83–95 (2003).
30. Krapp, A., Del Rosario, E. C. & Simanis, V. The role of *Schizosaccharomyces pombe dma1* in spore formation during meiosis. *J. Cell Sci.* **123**, 3284–3293 (2010).
31. Matos, J. *et al.* Dbf4-dependent CDC7 kinase links DNA replication to the segregation of homologous chromosomes in meiosis I. *Cell* **135**, 662–678 (2008).
32. Clyne, R. K. *et al.* Polo-like kinase Cdc5 promotes chiasmata formation and cosegregation of sister centromeres at meiosis I. *Nature Cell Biol.* **5**, 480–485 (2003).
33. Attner, M. A., Miller, M. P., Ee, L. S., Elkin, S. K. & Amon, A. Polo kinase Cdc5 is a central regulator of meiosis I. *Proc. Natl Acad. Sci. USA* **110**, 14278–14283 (2013).
34. Lee, B. H., Amon, A. & Prinz, S. Spo13 regulates cohesin cleavage. *Genes Dev.* **16**, 1672–1681 (2002).
35. Monje-Casas, F., Prabhu, V. R., Lee, B. H., Boselli, M. & Amon, A. Kinetochore orientation during meiosis is controlled by Aurora B and the monopolin complex. *Cell* **128**, 477–490 (2007).
36. Corbett, K. D. *et al.* The monopolin complex crosslinks kinetochore components to regulate chromosome–microtubule attachments. *Cell* **142**, 556–567 (2010).
37. Sarangapani, K. K. *et al.* Sister kinetochores are mechanically fused during meiosis I in yeast. *Science* **346**, 248–251 (2014).
38. Gregan, J. *et al.* The kinetochore proteins Pcs1 and Mde4 and heterochromatin are required to prevent merotelic orientation. *Curr. Biol.* **17**, 1190–1200 (2007).
39. Tada, K., Susumu, H., Sakuno, T. & Watanabe, Y. Condensin association with histone H2A shapes mitotic chromosomes. *Nature* **474**, 477–483 (2011).
40. Sakuno, T., Tada, K. & Watanabe, Y. Kinetochore geometry defined by cohesion within the centromere. *Nature* **458**, 852–858 (2009).
41. Tachibana-Konwalski, K. *et al.* Spindle assembly checkpoint of oocytes depends on a kinetochore structure determined by cohesin in meiosis I. *Curr. Biol.* **23**, 2534–2539 (2013).
42. Kagami, A. *et al.* Acetylation regulates monopolar attachment at multiple levels during meiosis I in fission yeast. *EMBO Rep.* **12**, 1189–1195 (2011).
43. Ahonen, L. J. *et al.* Polo-like kinase 1 creates the tension-sensing 3F3/2 phosphopeptide and modulates the association of spindle-checkpoint proteins at kinetochores. *Curr. Biol.* **15**, 1078–1089 (2005).
44. Kang, Y. H. *et al.* Self-regulated Plk1 recruitment to kinetochores by the Plk1–PBIP1 interaction is critical for proper chromosome segregation. *Mol. Cell* **24**, 409–422 (2006).
45. Hodges, C. A., Revenkova, E., Jessberger, R., Hassold, T. J. & Hunt, P. A. SMC1 $\beta$ -deficient female mice provide evidence that cohesins are a missing link in age-related nondisjunction. *Nature Genet.* **37**, 1351–1355 (2005).
46. Lister, L. M. *et al.* Age-related meiotic segregation errors in mammalian oocytes are preceded by depletion of cohesin and Sgo2. *Curr. Biol.* **20**, 1511–1521 (2010).
47. Chiang, T., Duncan, F. E., Schindler, K., Schultz, R. M. & Lampson, M. A. Evidence that weakened centromere cohesion is a leading cause of age-related aneuploidy in oocytes. *Curr. Biol.* **20**, 1522–1528 (2010).
48. Jessberger, R. Age-related aneuploidy through cohesion exhaustion. *EMBO Rep.* **13**, 539–546 (2012).
49. Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nature Rev. Genet.* **2**, 280–291 (2001).
50. Nagaoka, S. I., Hodges, C. A., Albertini, D. F. & Hunt, P. A. Oocyte-specific differences in cell-cycle control create an innate susceptibility to meiotic errors. *Curr. Biol.* **21**, 651–657 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank K. Tachibana-Konwalski for the C57BL/6 *Mlh1*<sup>tm1Liskay</sup> knockout mice and J. Ellenberg for a macro for automated microscopy. We also thank J. Lee, T. Hirano, S. Fujiyama, Y. Yamazumi, as well as the Gotoh laboratory for technical advice and all members of the Watanabe laboratory for their support and discussion. This work was supported in part by a JSPS Research Fellowship (to J.K.), SAF2011-25252 (to A.M.P.), a research grant from Uehara Memorial Foundation, a RIKEN CDB intramural grant and a Grant-in-Aid for Young Scientists (B) (to T.S.K.), a Grant-in-Aid for Scientific Research on Innovative Areas, a Grant-in-Aid for Scientific Research (C) (to K.I.), and a Grant-in-Aid for Specially Promoted Research (to Y.W.) from MEXT, Japan.

**Author Contributions** J.K., supported by K.I. performed most of the experiments in mice. K.I. and N.T. generated *Meikin* knockout mice. A.M.P. provided *Sgo2* knockout mice. A.N. isolated MEIKIN in yeast two-hybrid screening. B.A., S.Y., A.K., T.I. and T.S. performed experiments in fission yeast. Y.S. and T.S.K. performed live imaging. The experimental design and interpretation of data were conducted by J.K., K.I., B.A., S.Y., A.K., T.I., T.S.K., Y.T. and T.S. Y.W. supervised the project, and wrote the paper with input from all authors.

**Author Information** Sequence data are deposited with NCBI GenBank under accession numbers AB987828 for mMEIKIN and AB987829 for hMEIKIN. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.W. ([ywatana@iam.u-tokyo.ac.jp](mailto:ywatanab@iam.u-tokyo.ac.jp)).

## METHODS

**Yeast two-hybrid screening and assay.** For yeast two-hybrid screening, mouse *CENP-C* cDNA encoding the C terminus (amino acids 692–906) and mouse *Meikin* cDNA encoding the C terminus (amino acids 272–434) were subcloned into the vector pGBKT7. The bait strains were raised by transforming the bait vector into the yeast strain AH109. A matchmaker mouse testis cDNA library (Clontech) was transformed into the bait strains, and positive transformants were selected on nutrition-restricted plates (SD-trp-leu-his-ade, +10 mM 3AT). Positive transformants were further examined by blue/white assay to confirm the interaction. Prey plasmids were extracted from the candidate clones and sequenced. To exclude false positive clones, the candidate prey plasmids were retransformed into the yeast strain AH109 along with the bait vector or negative control bait vector of p53.

For yeast two-hybrid assays, the following cDNAs were subcloned into the bait vector pGBKT7: hCENP-C full-length, hCENP-C C-terminal region (amino acids 732–945), mPLK1 full length, hPLK1 full-length. The following cDNAs were subcloned into prey vector pACT2: MEIKIN full-length, MEIKIN exons1–12 (amino acids 1–384), MEIKIN exons 13–14 (amino acids 385–434). The following cDNAs were subcloned into the prey vector pGADT7: hMEIKIN N terminus (amino acids 1–264), hMEIKIN C terminus (amino acids 259–373). These bait and prey preparations were co-transformed into the yeast strain AH109.

**Mice.** *Sgo2* and *MLH1*<sup>tm1Liskay</sup> knockout mice were previously reported<sup>12,51</sup>. All single and double knockout mice were congenic with the C57BL/6 background. Whenever possible, each knockout animal was compared to littermates or age-matched non-littermates from the same colony, unless otherwise described. Animal experiments were approved by the Institutional Animal Care and Use Committee (approval numbers 23001, 23013, 24001, 2512 and 2608 at IMCB and AH23-05-04 at CDB). No statistical method was used to predetermine sample size.

**Generation of *Meikin* (493040A10rik) knockout mouse and genotyping.** The targeting vector was designed to disrupt exon 4 of the *Meikin* genomic locus. Targeting arms of 4.40 kb and 4.98 kb fragments, 5' and 3' to the exon 4 of *Meikin* gene respectively, were generated by PCR from mouse C57BL/6 BAC clone (RPC123-32113) and directionally cloned flanking pGK-Neo-polyA and DT-A cassettes. The homologous recombinant cells were isolated using Baltha1 ES cells derived from C57BL/6N<sup>52</sup>, and chimaeric mice were generated by morula injection (host ICR) of recombinant ES cells (Kumamoto University, CARD). The G418-resistant ES clones were screened for homologous recombination with the *Meikin* locus by PCR using primers: MEIKIN-7547F: 5'-GTTCAAGTTTACCTCCCGGTCTGAC and Neo3R: 5'-TACCGGTGGATGTGGAATGTGTGC for the left arm (4,734 bp), Neo104F: 5'-ggacgctatcaggatagcgtggc and MEIKIN-18860R: 5'-ACTCGCCACTGACTTCTCTGTGAGC for the right arm (5,531 bp).

Southern blotting was used to confirm correctly targeted ES cell clones. DNA was digested with Pvu II for Southern blots with an external 5'-probe. Chimaeric males were mated to C57BL/6 females and the progeny were genotyped by PCR using the following primers. Ex3F (common-forward): 5'-CCCCAGAGGAAAAGACACCACC-3', Ex4R (wild-type-reverse): 5'-CTCGACAACAAGCTGTCCATCTC-3' and Neo4R (mutant-reverse): 5'-CATGAGTGGGAGGAATGAGCTGGC-3' for the *Meikin* mutant allele (3,473 bp) and the wild-type allele (1,750 bp).

**Histological analyses.** The testes, epididymides or ovaries from mice (12-week-old males, 8-week-old females) were fixed in Bouin's solution and embedded in paraffin. Sections were prepared on APS-coated slides (Matsunami) at 5 µm thickness. The slides were dehydrated and stained with haematoxylin and eosin.

**Antibody production.** Polyclonal antibodies against mouse MEIKIN (amino acids 1–434) and MEIKIN C-terminal (amino acids 317–434) were generated by immunizing rabbits and mice. Polyclonal antibodies against human MEIKIN N- and C-terminal regions were generated by immunizing ICR mice. All His-tagged recombinant proteins were produced by inserting cDNA fragments in-frame with pET19b or pET28c (Novagen). All His-tagged recombinant proteins were purified by Ni-NTA (QIAGEN) under denaturing conditions using 6 M HCl-Guanidine. The antibodies were affinity-purified from the immunized serum with immobilized peptides on CNBr-activated Sepharose (GE healthcare).

**Antibodies.** The following primary antibodies were used for immunoblot (IB) and immunofluorescence (IF) studies: mouse anti-tubulin (IB, 1:5,000, IF, 1:1,000, DM1A, Sigma, catalogue no. T9026), human ACA (IF, 1:20, MBL, catalogue no. NA-8184), mouse anti-PLK1 (IB and IF, 1:1,000, Abcam, catalogue no. ab17056), rabbit anti-CENP-U/MLF1 phosphoT78 (IF, 1:100, Abcam, catalogue no. ab34911), rabbit anti-histone H3 (IB, 1:1,000, Abcam, catalogue no. ab1791), rabbit anti-GFP (IF, 1:1,000, Invitrogen, catalogue no. A11122). The following polyclonal antibodies were described previously: rabbit anti-mCENP-C (IB and IF, 1:1,000), rabbit anti-mREC8 (IF, 1:500), mouse anti-mREC8 (IF, 1:500), mouse anti-mSYCP3 (IF, 1:1,000)<sup>53</sup>, Mouse anti-mSGO2 (IF, 1:500)<sup>54</sup>, Rabbit anti-mSGO2 (IF, 1:1,000)<sup>51</sup>, Rat anti-mSYCP3 polyclonal antibody (IF, 1:1,000)<sup>55</sup>.

The following secondary antibodies were used for immunofluorescence studies: Alexa488-conjugated donkey anti-mouse (Invitrogen, catalogue no. A21202), donkey anti-rabbit (Invitrogen, catalogue no. A21206), Alexa555-conjugated goat anti-mouse (Invitrogen, catalogue no. A21422), donkey anti-rabbit (Invitrogen, catalogue no. A31572), goat anti-human (Invitrogen, catalogue no. A21433), goat anti-rat (Invitrogen, catalogue no. A21434), Alexa647-conjugated donkey anti-mouse (Invitrogen, catalogue no. A31571), goat anti-rabbit (Invitrogen, catalogue no. A21244), goat anti-human (Invitrogen, catalogue no. A21445), goat anti-rat (Invitrogen, catalogue no. A21247).

**In vitro oocyte culture.** Ovaries collected from 6-week-old to 12-week-old female mice were used for this study after 46 to 48 h of treatment with 5 IU of pregnant mare serum gonadotropin. GV oocytes were isolated by puncturing the follicles in M2 medium (Sigma) containing 250 µM 3-isobutyl-1-methyl-xanthine (IBMX, Sigma) to maintain prophase arrest. To induce resumption of meiosis, the oocytes were cultured in M16 medium (Sigma) supplemented with 10% FBS in a 5% CO<sub>2</sub> atmosphere at 37 °C. Oocytes that had not undergone GV breakdown (GVBD) by 90 min were removed from the experiment. For oocyte drug treatment, 100 nM BI 2536 (ChemieTek), 5 µM reversine (Sigma) and 10 µM nocodazole (Sigma) were added to the culture medium at the indicated time point; control oocytes were treated with an equivalent volume of DMSO.

**Immunofluorescence microscopy of spermatocytes and oocytes.** Squashed spermatocytes were prepared from adult male mice as described previously<sup>12</sup> with modification. Briefly, seminiferous tubules were minced, and fixed in fixation buffer of 2% paraformaldehyde (PFA)/0.1% Triton X-100/PBS. The cell suspension was filtered through a cell strainer (BD Falcon) to remove debris, pipetted repeatedly and centrifuged. The cell pellets were suspended in the fixation buffer and gently squashed by the cover glass. After fixation, the slides were frozen in liquid nitrogen. For immunostaining, the frozen slides were immersed in PBS and the coverslips were removed.

Fixed whole mount oocytes was prepared as described previously<sup>11,25</sup> with minor modification. After *in vitro* culture, the oocytes were fixed in the fixation buffer for 30 min at room temperature, washed with PBS and then blocked with 3% BSA/PBT (0.1% Triton X-100/PBS) for 1 h at room temperature. For measurement of sister kinetochore distance at early prometaphase I, oocytes were exposed to 1% Pronase (Sigma) to remove the zona pellucida, and fixed in the fixation buffer on glass slides for 3 h.

Oocyte chromosome spreads were prepared as described previously<sup>56</sup>. Briefly, after zona pellucida removal, oocytes were transferred onto glass slides and fixed in a solution of 1% PFA/0.15% Triton X-100/distilled H<sub>2</sub>O adjusted to pH 9.2. After a quick dry, chromosomes were immunostained as described above. Fetal oocyte chromosome spreads were prepared as described previously<sup>57</sup>.

For immunostaining spermatocytes and oocytes, samples were briefly washed in PBS, blocked with 3% BSA/PBS for 10 min at room temperature and incubated with primary antibodies in 3% BSA/PBS for 1 h and secondary antibodies for 1 h at room temperature. In whole mount oocytes, incubation was performed at 4 °C for overnight for each antibody. The slides were washed with PBS, and mounted using VECTASHIELD medium with DAPI (Vector Laboratories).

Images were acquired on an IX-70 microscope (Olympus) equipped with a CoolSnap HQ CCD camera (Roper Scientific), DeltaVision Core system (GE Healthcare). Only whole mount oocytes images were captured with an FV1000 confocal laser scanning microscope at 1 µm intervals and processed with FLUOVIEW software (Olympus).

The projection of the images, the quantification of signal intensity and the measurement of sister kinetochore distance were carried out with the SoftWorx software program (GE Healthcare). For the measurement of sister kinetochore distance, images were acquired with z-sections encompassing the entire nuclei. The peak-to-peak distance of CENP-C signals was measured for a pair of sister kinetochores in structurally preserved nuclei. The sister kinetochores distance was measured by calculating the square root of  $x^2 + y^2 + z^2$ .

**Exogenous expression of GFP-tagged MEIKIN variants in testis.** For exogenous expression of GFP-tagged MEIKIN variants in testis, plasmid DNA was injected into live-mouse testes as described previously<sup>58</sup>. *Meikin* variants were cloned in pCAG vector. Plasmid DNA was injected into live-mouse testes as described previously<sup>58</sup>. Briefly, male mice at 16–20 days postpartum were anaesthetized with pentobarbital and the testes were pulled from the abdominal cavity. Then 50 µg of plasmid DNA (10 µl of 5 µg µl<sup>-1</sup> DNA solution) was injected into each testis using glass capillaries under a stereomicroscope (M165C; Leica). Testes were held between a pair of tweezer-type electrodes (CUY21; BEX), and electric pulses were applied four times and again four times in the reverse direction at 30 V for 50 ms for each pulse. The testis was then returned to the abdominal cavity, and the abdominal wall and skin were closed with sutures. Immunostaining was performed 24–48 h after electroporation to assess the localization of the GFP-tagged *Meikin* variants in spermatocyte using rabbit anti-GFP (Invitrogen). Because the efficiency of GFP-tagged



protein expression in meiotic prophase is 5–10%, we detected kinetochore localization of GFP-tagged MEIKIN from GFP-positive cells.

**Live confocal imaging of oocytes.** Oocytes were microinjected with *in vitro* transcribed RNAs encoding 2mEGFP–CENP-C (0.8 pg) and H2B–mCherry (0.2 pg), and incubated for 3 h before IBMX washout. Timing of GVBD was determined by low resolution time-lapse imaging. Oocytes not undergoing GVBD within 1.5 h after the IBMX washout were removed from further imaging analysis. High resolution imaging was started around 5.5 h after GVBD. Imaging was performed with Zeiss LSM710 equipped with a 40× C-Apochromat 1.2 W Corr M27 objective lens (Carl Zeiss) and a 3D multi-location tracking macro<sup>59</sup>. We acquired 17 z axis confocal sections (every 1.5 μm) of 2-time-averaged 512 × 512 pixel xy images, which covered a volume of 30.36 μm × 30.36 μm × 25.5 μm at 5-min time intervals. 2mEGFP–CENP-C signals were peak-enhanced and background-subtracted as previously described<sup>25</sup>. The images were 3D-reconstructed with Imaris (Bitplane) and the kinetochore signals during anaphase were manually tracked. The kinetochore tracks were visualized by POV-Ray (<http://www.povray.org>).

**PCR with reverse transcription.** Total RNA was isolated from tissues using TRIzol (Invitrogen). cDNA was generated from 0.5 μg of total RNA using Superscript III (Invitrogen) followed by PCR amplification using Ex-Taq polymerase (Takara) and template cDNA (derived from 1 ng RNA equivalent). Sequences of primers used to generate RT-PCR products from cDNA are as follows: Meikin-F2:5'-agatgga cagctgtgtcagta-3'; Meikin k-R2:5'-ctcagcaataacacctcagaagc-3'; GAPDH-F:5'-ttaccaccatggagaagc-3'; GAPDH-R:5'-ggcatggactgtgtcagta-3'; mSMC1-3516F: 5'-ttatcatcaaggacagctcaacttcg-3'; mSMC1-3702R: 5'-ctattgttcttgggttgggttg-3'; Zfp438-F1:5'-gtatgcaaggacgtgacactac-3'; Zfp438-R1:5'-cattgtctctctctgtgta-3'; BC051142-F1:5'-ccttatcaacctgtcagctctt-3'; BC051142-R1: 5'-ctcgaaatcttctggaga gcagta-3'; Nsun7-F1:5'-cctctcgattaccatattctgcc-3'; Nsun7-R1: 5'-tatcaagccttcac agatgtggac-3'; Spesp1-F1:5'-catgaagctgtgtgtcctagtgtg-3'; Spesp1-R1:5'-gatggaccaga atgtgtactttg-3'; Snf1lk-F1:5'-gaggtccagctcatgaactttt-3'; Snf1lk-R1:5'-gctagcttg atatcattgtgtg-3'; Acbd3-F1:5'-caggagcagcactatcagcagta-3'; Acbd3-R1:5'-cgact tctctcagctactgtaa-3'; Ncor1-F1:5'-aactcttctgtggagtgactct-3'; Ncor1-R1:5'-tgcc caggaataggagattagac-3'.

**Immunoprecipitation and mass spectrometry using testis extracts.** Testis chromatin-bound and -unbound extracts and immunoprecipitation were prepared as described previously<sup>53</sup>. Briefly, immunoprecipitations were performed with protein A–Dynabeads (Invitrogen)-conjugated rabbit anti-mCENP-C, rabbit anti-MEIKIN and control rabbit IgG (5 mg equivalent) from the chromatin-bound fraction prepared from 40–60 wild-type testes (3-week-old). Co-immunoprecipitated proteins were run in 4–12% NuPAGE gels (Invitrogen) in MOPS-SDS buffer and immunoblotted or analysed by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS).

Fractions containing the MEIKIN immunoprecipitates were concentrated by precipitation with 10% trichloroacetic acid. The derived precipitates were dissolved in 7 M urea, 50 mM Tris–HCl (pH 8.0), 5 mM EDTA solution, with 5 mM DTT at 37 °C for 30 min, and cysteine SH groups were alkylated with 10 mM iodoacetamide at 37 °C for 1 h. After alkylation, the solutions were desalted by methanol/chloroform precipitation, and the precipitates were dissolved in 2 M urea, 50 mM Tris–HCl buffer and subjected to trypsin gold (Promega) digestion overnight at 37 °C. The resulting mixture of peptides was applied directly to the LC–MS–MS analysis system (Zaplous, AMR, Tokyo, Japan) using Finnigan LTQ mass spectrometry (Thermo Scientific) and a reverse phase C18 ESI column (0.2 × 50 mm, LC assist). The protein annotation data were verified in the mouse NCBI sequences using Bioworks software (ver. 3.3; Thermo Scientific) with quantitation featuring the SEQUEST search algorithm.

**Cloning of human Meikin homologue cDNA.** Human Meikin homologue cDNAs encoding the N terminus (amino acids 1–264) and C terminus (amino acids 259–373) were cloned from a human testis cDNA library (Takara) using the following primer sets: NheI–hMEIKIN–N-1F: tACCGGT gctagc atgtggccgtacgg gctatacc and NotI–hMEIKIN–N-795R: tACtCGGTgctgCgCttattctgtcaactt gctttttt for N terminus, NheI–hMEIKIN–C-1F: tACCGGTgctagcagcagaacaaa ttccagtact and NotI–hMEIKIN–C-354R: tACtCGGTgctgCgCtcatgccattttatt gatacttg for C terminus.

**Immunostaining of human seminiferous tubule sections.** Frozen sections of adult human testis were purchased from Biochain. After washing with PBS/0.1% Triton X100, immunostaining was performed as described above.

**Schizosaccharomyces pombe strain.** All strains used in this study are described in Supplementary Table 1. Complete medium (YE), minimal medium (s.d. and MM) and sporulation-inducing medium (SPA) were used for culture. The construction of *moa1Δ*, *GFP-3Pk-moa1*<sup>+</sup> and *rec12Δ* were described previously<sup>23</sup>. The C-terminal tagging of endogenous *plol*<sup>+</sup> by GFP was performed using the PCR-based gene targeting method for *S. pombe*<sup>60</sup>. To construct 3HA or GFP tagged *Plol* at its N terminus, the coding sequence of 3HA or GFP flanked by the 5'- and ORF containing 3' untranslated region (UTR) of the *plol* gene was

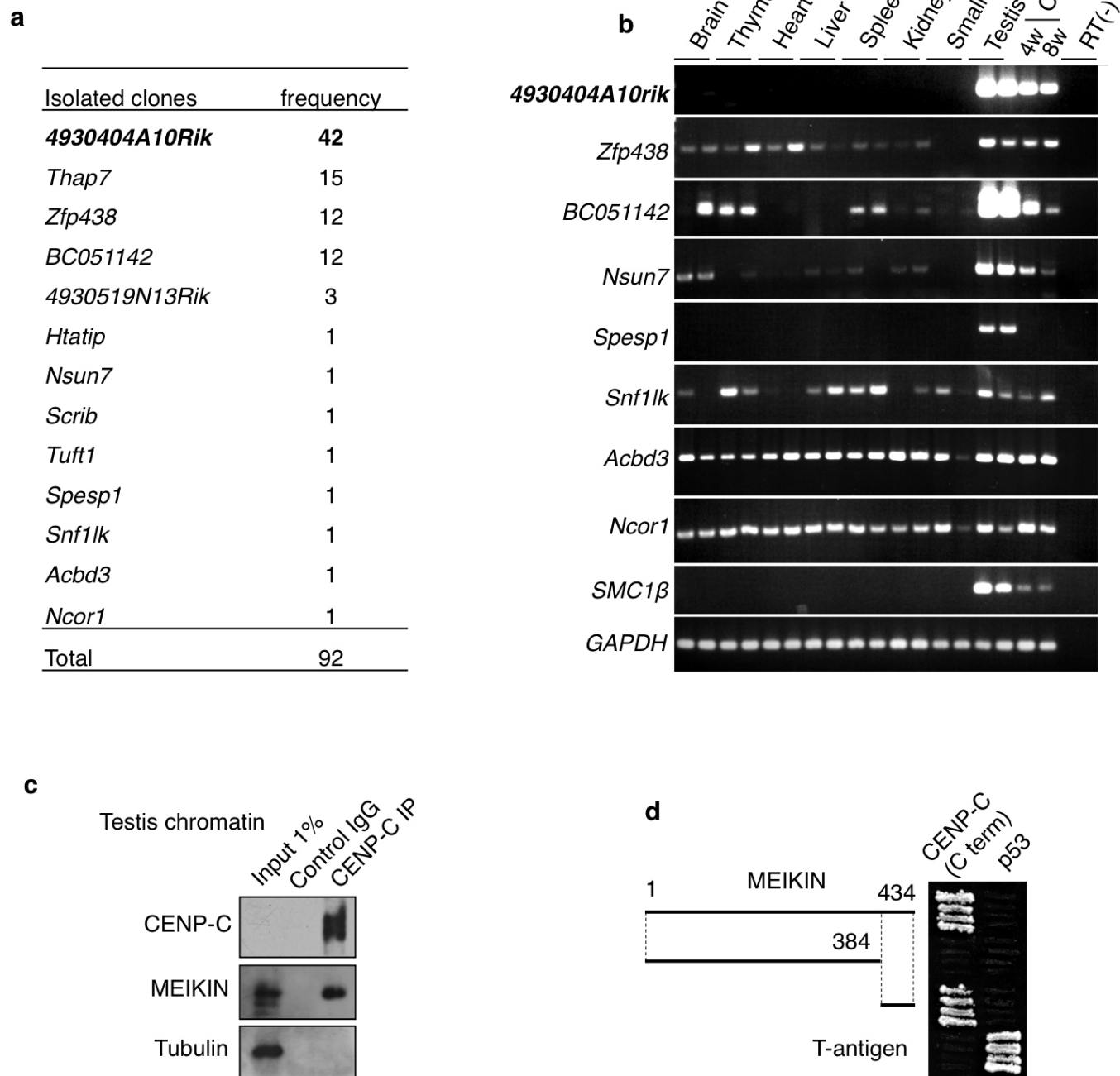
cloned into pUC119 (3HA; vp183 and GFP; pak64). The 5'-UTR-3HA/GFP-*plol*-3'-UTR fragment amplified by PCR was transformed into temperature-sensitive *plol1-ts2::ura4*<sup>+</sup> strain (from Ian Hagan). To construct the *plol1-tev* allele, the recognition sequence for TEV protease Glu-Asp-Leu-Tyr-Phe-Gln-Gly(-Ala-Ser) was inserted at both the 341 amino acid and 405 amino acid sites in pak64 or 5'-UTR-*plol*-3'-UTR (untagged version) fragment cloned into pUC119 by site-directed mutagenesis. The PCR-amplified fragment contacting the *plol1-tev* allele was introduced into *plol1-ts2::ura4* strain. The construction of Cnp3C-CFP-TEV was described previously<sup>23</sup>, while the promoter was exchanged from *Padh81* to meiosis-specific *Prec8*. The T101A mutation of *moa1* was introduced into 5'-UTR-3Pk-*moa1*-3'-UTR fragment cloned in pBluescript by using the KpnI and SacI sites by site-directed mutagenesis. The KpnI and SacI-digested fragment containing 5'-UTR-3Pk-*moa1*-T101A-3'-UTR were introduced into *moa1::ura4*<sup>+</sup> strain. To visualize tubulin, besides the pREP81-CFP-*atb2*<sup>+</sup>, *Padh13* (a weaker promoter version of the *Padh1*)-mCherry-*atb2*<sup>+</sup> integrated at the Z locus using the *nat*<sup>r</sup> marker was used<sup>61</sup>. To localize the Spo13 protein to kinetochores in fission yeast meiotic cells, the SPO13 ORF cloned from *S. cerevisiae* W303 genome was fused to the C terminus of Cnp3C-CFP, and expressed by meiosis-specific *Pspo5del* (authentic TATA sequence of *Pspo5* was deleted to reduce the expression level, but it still shows a stronger expression level than that of *Prec8*). The resulting plasmid (pHBCS5del-Cnp3C-CFP-SPO13) was linearized and integrated at the C locus<sup>61</sup> using the *hyg*<sup>r</sup> marker. To construct *spo13-5A*, five STSP residues (amino acids 132 to 136) of SPO13 gene in pHBCS5del-Cnp3C-CFP-SPO13 were substituted to alanines by site directed mutagenesis. To target *Moa1* or the *Moa1*-T101A protein to kinetochores in fission yeast meiotic cells, *moa1*<sup>+</sup> or *moa1-T101A* fragment were cloned into pHBCS5del-Cnp3C-CFP, and integrated at the C locus using the *hyg*<sup>r</sup> marker.

**Synchronization of fission yeast meiotic cells.** For microscopic observation of *cen2*-GFP or GFP-*Plol*, logarithmically growing cells were collected and suspended in 20 mg ml<sup>-1</sup> leucine and spotted on SPA and incubated at 26.5 °C for ~12–15 h (*cen2*-GFP observation in spored asci) or ~6–8 h (GFP-*Plol* observation at metaphase I zygote with short spindle). Fluorescence images were taken using a microscope (Axioplan2, Zeiss) equipped with a cooled CCD camera (Quantix, Photometrics) and AxioVision software (Zeiss). Seven z sections (0.45 μm each) for GFP signals were converted into single two-dimensional images by taking the maximum signal at each pixel position in the images. For immunoprecipitation analysis of Pk-Moa1, we used haploid cells containing the *pat1-114* allele. Cells were grown in MM liquid medium including NH<sub>4</sub>Cl (MM+N) to a density of 5 × 10<sup>6</sup> cells ml<sup>-1</sup> at 25 °C, then resuspended in MM medium lacking NH<sub>4</sub>Cl (MM-N) at a density of 1 × 10<sup>7</sup> cells ml<sup>-1</sup> at 25 °C for 15 h. To induce meiosis, cells were incubated at 34 °C, and cells were collected 4 h after meiosis induction, the period of meiotic prophase<sup>23</sup>.

**Immunoprecipitation from yeast extracts.** Before harvesting the meiotic cells, 1 mM PMSF (phenylmethylsulfonyl fluoride, Sigma) was added to reduce protein degradation. Cell extracts (~8 × 10<sup>7</sup> cells) were prepared by beating the cells with beads in HB IP buffer (25 mM MOPS, 5 mM EGTA, 15 mM MgCl<sub>2</sub>, 150 mM KCl, 50 mM β-glycerophosphate, 15 mM *p*-nitrophenylphosphate, 1 mM DTT, 0.1 mM sodium vanadate, 0.8% NP40, 1 mM PMSF). Immunoprecipitation was performed by incubating the whole cell extracts for 1 h at 4 °C with anti-Pk monoclonal antibody (Serotec), and 20 μl of protein A-Sepharose (GE-healthcare) for 2 h at 4 °C. Whole cell extracts and immunoprecipitates were subjected to immunoblot analysis using mouse anti-HA monoclonal antibody 12CA5 (Roche) or anti-Pk monoclonal antibody.

- Baker, S. M. *et al.* Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. *Nature Genet.* **13**, 336–342 (1996).
- Kim-Kaneyama, J. R. *et al.* Hic-5 deficiency enhances mechanosensitive apoptosis and modulates vascular remodeling. *J. Mol. Cell. Cardiol.* **50**, 77–86 (2011).
- Ishiguro, K., Kim, J., Fujiyama-Nakamura, S., Kato, S. & Watanabe, Y. A new meiosis-specific cohesin complex implicated in the cohesin code for homologous pairing. *EMBO Rep.* **12**, 267–275 (2011).
- Kawashima, S. A., Yamagishi, Y., Honda, T., Ishiguro, K. & Watanabe, Y. Phosphorylation of H2A by Bub1 prevents chromosomal instability through localizing shugoshin. *Science* **327**, 172–177 (2010).
- Morimoto, A. *et al.* A conserved KASH domain protein associates with telomeres, SUN1, and dynactin during mammalian meiosis. *J. Cell Biol.* **198**, 165–172 (2012).
- Chambon, J. P., Hached, K. & Wassmann, K. Chromosome spreads with centromere staining in mouse oocytes. *Methods Mol. Biol.* **957**, 203–212 (2013).
- Hodges, C. A. & Hunt, P. A. Simultaneous analysis of chromosomes and chromosome-associated proteins in mammalian oocytes and embryos. *Chromosoma* **111**, 165–169 (2002).
- Shibuya, H., Ishiguro, K. & Watanabe, Y. The TRF1-binding protein TERB1 promotes chromosome movement and telomere rigidity in meiosis. *Nature Cell Biol.* **16**, 145–156 (2014).
- Rabut, G. & Ellenberg, J. Automatic real-time three-dimensional cell tracking by fluorescence microscopy. *J. Microsc.* **216**, 131–137 (2004).

60. Bähler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943–951 (1998).
61. Sakuno, T., Tanaka, K., Hauf, S. & Watanabe, Y. Repositioning of Aurora B promoted by chiasmata ensures sister chromatid mono-orientation in meiosis I. *Dev. Cell* **21**, 534–545 (2011).



**Extended Data Figure 1 | MEIKIN (4930404A10rik) was identified as a meiosis-specific CENP-C binding protein.** **a**, Yeast two-hybrid screening was performed using CENP-C C terminus (amino acids 692–906) as bait and a mouse testis cDNA library as prey. A total of  $11.95 \times 10^7$  colonies were screened on selective (SD-Trp-Leu-His-Ade, +10 mM 3AT) plates using the AH109 tester strain. The number of clones isolated by screening is summarized. **b**, To search for a meiosis-specific candidate from the isolated clones, tissue-specific expression patterns of CENP-C interactors were examined by RT-PCR. RNA was extracted from each tissue of both males and females. Testis

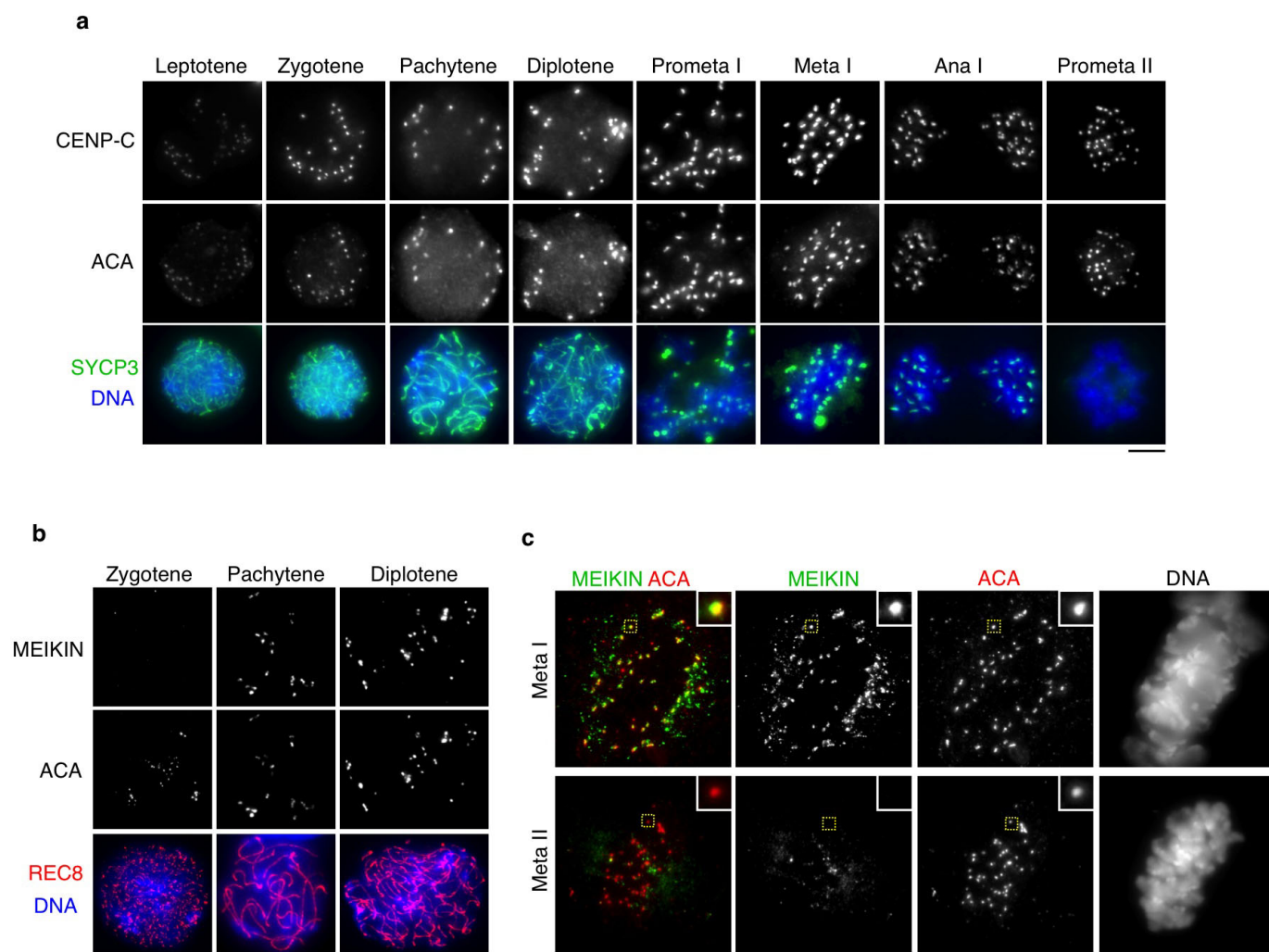
RNA was derived from 8-week-old males. Ovary RNA was derived from 4- and 8-week-old females. RT(–) indicates control PCR without reverse transcription. Note that the expression of *4930404A10rik* is restricted to testis and ovary, as is that of *SMC1β*. **c**, Immunoprecipitates from mouse testis chromatin extracts using anti-CENP-C antibody or control IgG were analysed by the indicated antibody. **d**, The C-terminal domain of MEIKIN (4930404A10rik) (amino acids 385–434) interacts with the CENP-C C terminus in yeast two-hybrid assay.



<i>M. musculus</i>	1	MDKIWHMGPGRGDYTRKKRAGERLNLTPKPDIALPGRTEALPGLKG-----KGKEOGLR
<i>R. norvegicus</i>	1	MDKIGRVRLRRVYIRKKRAGERLNLTPKPNLALPGKTP--LGLKG-----KGKEOGLQ
<i>H. sapiens</i>	1	-----MWPLRVYTRKKRECORLNLTPDGLGSPAKAEAPPFSKR-----KGKVHGLS
<i>C. lupus</i>	1	-----MWSLQVYTRKKRAGORLNLTPDQGPATAEPPVPVPGRRHSARRLKGKMHRLQ
<i>X. tropicalis</i>	1	-----MFGARLRALKRVGRSSPDPRADTQTASSVGRRAGDWK-----ATQKHRVQ
<i>G. gallus</i>	1	-----
<i>M. musculus</i>	54	KITEKKELSRLTGSSSQ-RPSLLSVTGGEHLQENSPGQETPEEKTTPC-ETVDTTFEMDS
<i>R. norvegicus</i>	52	KITEKEELSRLPTGSSSR-RPSQLSIPGGESLQENSAGKETPEEKTTPC-SSVTDAFETDS
<i>H. sapiens</i>	48	KIAEKAERSROGGSGSGPFSRPLGVTGKESLQENRSEDQDEKIASLRESVTDDLQVDS
<i>C. lupus</i>	55	KISEKAECSSGRGGGGPRPRSTPLRVGTGQKSLQENITNEEPADEKTAPLSESVPDDLQVDS
<i>X. tropicalis</i>	46	KRKPFKFKTMGKIKENIEASKHHDALDGSTSETAGSIHLKQIKGREKIHSESSSLSQEI
<i>G. gallus</i>	1	-----MTLPTGVSFAFLLECLDADS
<i>M. musculus</i>	112	LLSSTELVSGPAEQDDFSSCLPSCSNAELHTEST--DERGSSFPSPPELFRGSDCLD---W
<i>R. norvegicus</i>	110	FSSSAELVSGPSMQDDVSSCLPSCSNAESYTKNT--EDSGSSFPSPPELFRGADCLD---W
<i>H. sapiens</i>	108	SSSNSELVSGLSLHHGMASLLSYSVTDSYAEYKSFEES--FPSPPELFRKSDYLD---W
<i>C. lupus</i>	115	SSSNSELISGLTSQH-ISSSFPSYSLTDSYTEYKSFEESLSSFPSPPELFRGSDYLD---W
<i>X. tropicalis</i>	106	QQETDNTASTRYSETTSGITLPTLYQVSSSESETSILSTSVSSCLSPPEVLRGADSFDESEW
<i>G. gallus</i>	20	SAYGDTVAT-----DTTESFPSPETLRLDEECSCG---T
<i>M. musculus</i>	167	EHPKLEDYMFYKNSTLLDTSKAVVIEKAPQFANLSAVLSSSSKNYEKRHRKIGMTLAAQH
<i>R. norvegicus</i>	165	EHPKLEDYMLYKNSTLLDTSKAVVIEKAPQFANLSAVLSSSENIEKCRKIGMTLEAOY
<i>H. sapiens</i>	162	ECPNLEEHEMOWKNSTLLDTSKAVAIEKAPQFSNVSAIFSTSSSEDYQKCHRKTVMTVADQN
<i>C. lupus</i>	171	ECPKLEEHEMRCKNSTLLDTSKAVAIEKVPOFSNLSAILGTSSGDSQKCHREIATMLADQN
<i>X. tropicalis</i>	166	KSIFHNENLFCNSTLLETSAVNIDFLPEPTDVSNIFEPIDRVHLHGAKSNKNVKSNDNI
<i>G. gallus</i>	49	CNPDEFDFVKCKNSTLLDCSKAVAIDKILQISNISPILEPVLEDCKDQHIKRRKPKCNYS
<i>M. musculus</i>	227	LSPEPKYASNLASVVDNAASEVVF AEKTGPPTTEKTQKKPENES EDSGPLVQTKLSSGHP
<i>R. norvegicus</i>	225	ISPEPKYTSNLASVVP SAASGVICTEKTGPPTIKTTKKKPEKES EDGGPLVQTKLSSGHP
<i>H. sapiens</i>	222	VSPKAKCAS--NSEDNAACEILLAEKTCPSPTPEKTKKK-----
<i>C. lupus</i>	231	ISPKPKNTS--HSEDNAACEVLLAEKTYPSAPEKTKKKPEKGP EPRDTNFQTKLSFHHL
<i>X. tropicalis</i>	226	FEKQHVSTI---VSGKKIFRITREKEKTPRS DGGSDSKCHTRPEKKVKESCDTTDKNVK
<i>G. gallus</i>	109	SSELSVSTT---LAGKKICKITTAARETPALKSGTRCPSP-----
<i>M. musculus</i>	287	DNKALCSPLSSALESTAVR---YTLLPQPLEPV LKKGCILPDKQSKALLTSTPSSDIAEF
<i>R. norvegicus</i>	285	DSKVLCSPLSSAPESSAVR---DALLPQRLPALKKSCILPDNOSKVLLTSTPSSDIAEF
<i>H. sapiens</i>	259	-----KTNSSTPGKKNRGLLTSTPSSSETAGF
<i>C. lupus</i>	289	KIKIPSSHQRSVIETNAGRSITKVPFSQPLEPALRTNSSTPGKKS RGLLTSTPSSKTAGL
<i>X. tropicalis</i>	282	CKKRVSFGTVQTTEIPPNE-----YVDHCAGKNIDSSAKTDLLEQQKLSLSLTH
<i>G. gallus</i>	146	-----VGAGRKPDDEQTAEPKRLKCIKKEELS
<i>M. musculus</i>	344	VIDLSPVQNVSFEE LFPNVSN--YVNSSEIVPVSSL--QESSSNEFSPNTSEICCIIRSS
<i>R. norvegicus</i>	342	VIDLSPVQNVSFEE LFPNVSN--YVNSSEIVPVSSW--QESSLNEVSPNTSEICCIIRAS
<i>H. sapiens</i>	285	VIDLSSVQKASFEE LFPNVSN--YVNSNEIVPVSSL--QENSSNEFPANASEICCIIRTS
<i>C. lupus</i>	349	VIDLSSVQKASFEE LFPNVSN--YVNSNEIVPVSNL--QENSSNEFSPNTSEMCCIIRAT
<i>X. tropicalis</i>	331	KLDLSPVCVSSMEEEIIPNASGPFVHSIQIAPSSTLSMOKES EKQVLHEAKEICCIKAP
<i>G. gallus</i>	172	NLLEGNASSCAQLESAPANTS--AKSVKVTTEVLPS--RQTGDVVVTQCYRKEICSIVRTS
<i>M. musculus</i>	400	PGTRQMRRKDPAVKN-RCSPPKDVPLDIIMKTNGRT-----
<i>R. norvegicus</i>	398	PGTRQMRRKDATVKN-RCSPVPTDVPLDIIMKTNGRT-----
<i>H. sapiens</i>	341	PGTRQVKNGKVIVK KKKYSLPKDTPODIIIKMA-----
<i>C. lupus</i>	405	PGTRQVKSKDVIVK KKKYSPPKDIPODIIIKTNGRM-----
<i>X. tropicalis</i>	391	VCRCTSAILGVPIANPRLPPKHKIPKDIILTF-----
<i>G. gallus</i>	228	PGHRPSRHRQLPVRTKAFCLPEGVPEVDVITNSKTWICCKHR

**Extended Data Figure 2 | Sequence alignment of MEIKIN homologues in vertebrates.** Amino acid sequences of *M. musculus* 4930404A10Rik (NP\_083381), *R. norvegicus* (XP\_573090), *C. lupus* (XP\_003639413), *X. tropicalis* (XP\_002934413) and *G. gallus* (XP\_001234011) are derived from the NCBI protein database. *H. sapiens* data are derived from cDNA clones. 4930404A10Rik protein is conserved among vertebrates but it does not have

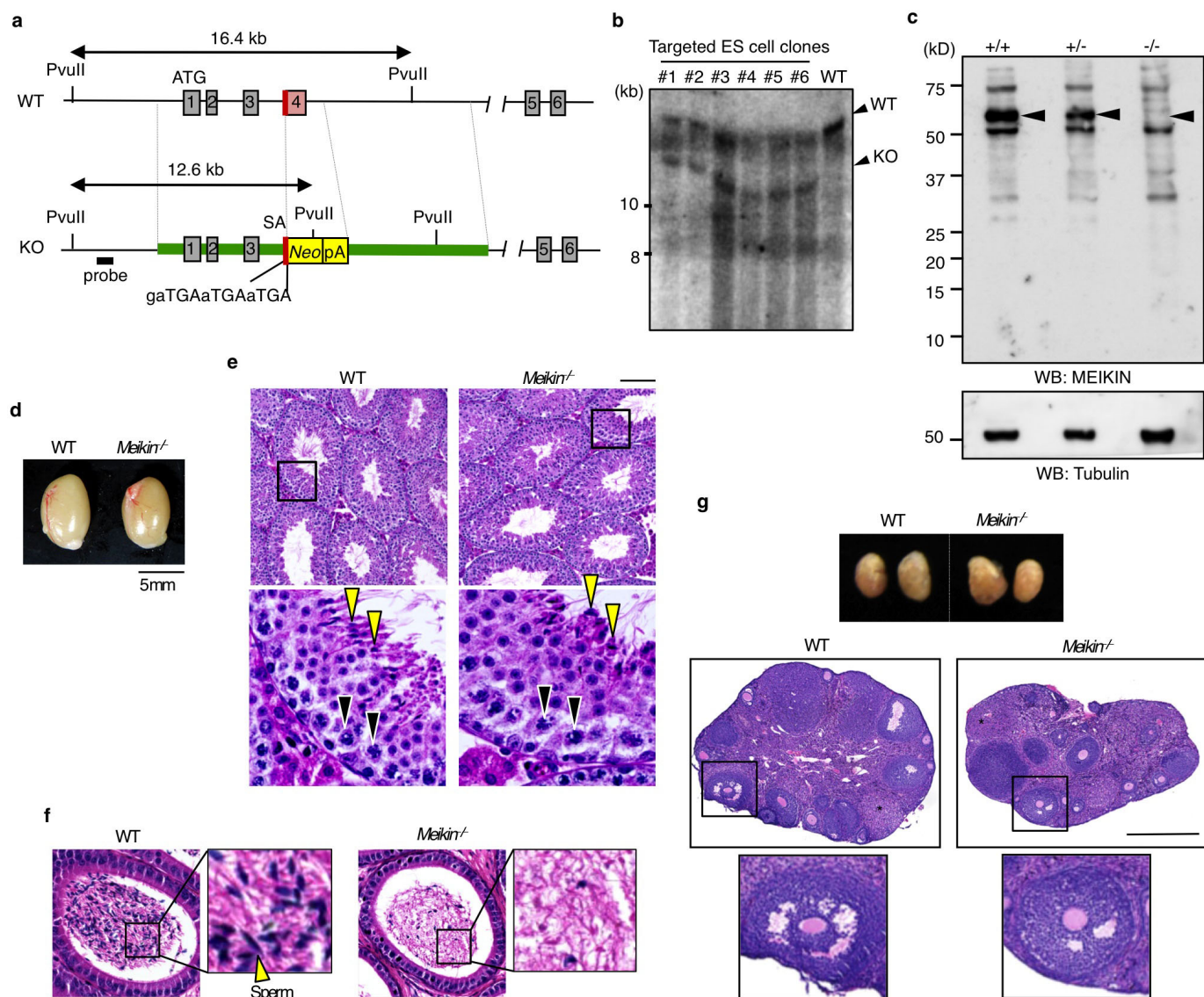
any known motif except for the polo-box binding motif (blue line) in mammalian proteins. Two-hybrid assays indicate that this motif of mouse MEIKIN is important for PLK1 binding (data not shown), although the motif is apparently not conserved in *Xenopus* and chicken. The C-terminal sequences (red lines) are required for the kinetochore localization (see Fig. 1c).



**Extended Data Figure 3 | Co-localization of CENP-C and ACA in spermatocytes and MEKIN localization in oocytes.** **a**, Squashed spermatocytes from wild-type were immunostained for CENP-C, ACA, SYCP3 and DAPI at the indicated stages during meiosis. CENP-C and ACA signals accumulate and co-localize at centromeres after zygotene throughout meiosis.

**b**, Chromosome spreads of oocytes from wild-type were immunostained for MEIKIN, ACA and REC8 at different meiotic stages. Zygotene from E15.5, pachytene and diplotene from E18.5 mice. **c**, Chromosome spreads of oocytes at metaphase I (5 h after GVBD) and metaphase II (16 h post GVBD) were stained for MEIKIN, ACA and DAPI. Scale bars, 5  $\mu$ m.



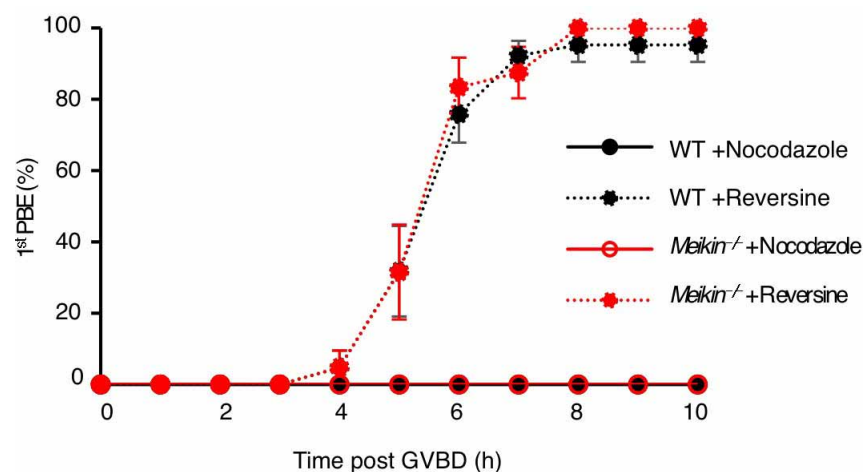


#### Extended Data Figure 4 | Generation of *Meikin*-knockout mice.

**a**, Schematic illustrations of the wild-type allele and targeted *Meikin*<sup>-/-</sup> allele are shown. Grey boxes represent exons. The targeted exon 4 contains the intact splicing acceptor (SA) sequence followed by a premature stop codon, resulting in disruption of the *Meikin* allele. Black bar probe for Southern blot. **b**, Southern blot of genomic DNA from wild type (+/+) and *Meikin* heterozygous (+/-) ES cells after Pvu II digestion. ES cell clone number 2 used to generate the mice. **c**, Immunoblot analysis of testis extracts prepared from mice with the indicated genotypes (4-week-old). In *Meikin*<sup>-/-</sup>, the specific bands probed by the anti-MEIKIN antibody are absent (shown as arrowheads).  $\alpha$ -tubulin is a loading control. **d**, Testes from 12-week-old wild-type (+/+) and

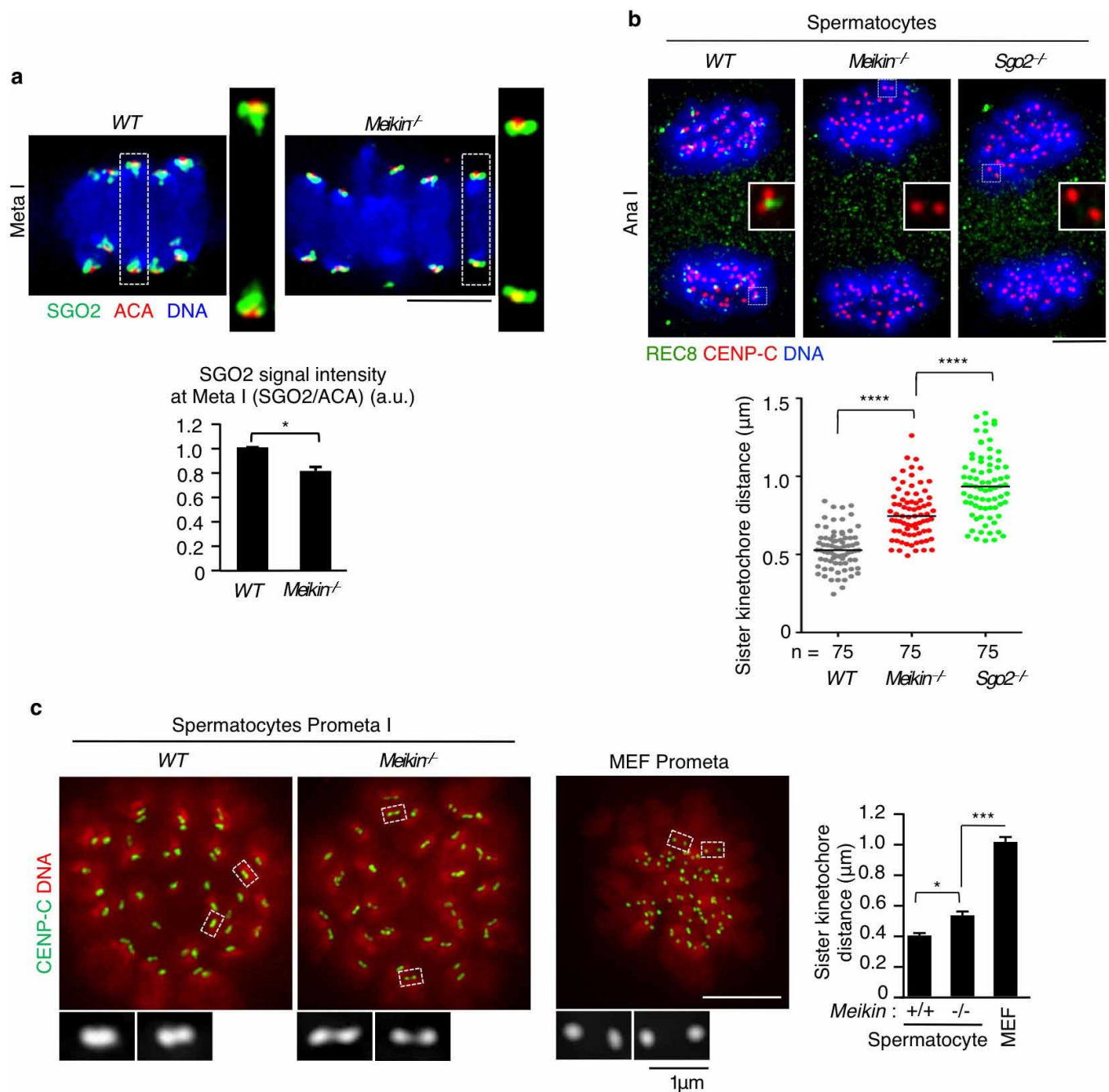
*Meikin*<sup>-/-</sup> (-/-) mice. **e**, Haematoxylin and eosin staining of a section of the testis (12-week-old) showed seminiferous tubules. Enlarged pictures of seminiferous tubules showed spermatocyte (black arrowheads) and spermatids (yellow arrowheads) in wild-type and *Meikin*<sup>-/-</sup>. Scale bar, 100  $\mu$ m. **f**, Haematoxylin-eosin staining of epididymis from 12-week-old mice shows reduced number of sperms in *Meikin*<sup>-/-</sup>. Enlarged images of sperms are shown. Scale bar, 50  $\mu$ m. **g**, A pair of ovaries (8-week-old) from the indicated genotypes (top). Haematoxylin and eosin-stained paraffin sections of ovaries from 8-week-old wild-type and *Meikin*<sup>-/-</sup> mice (middle). The antral-stage follicles with oocyte nuclei are magnified (bottom). Asterisks indicate corpora lutea. Scale bar, 500  $\mu$ m.





**Extended Data Figure 5 | The delay of anaphase I onset in *Meikin*<sup>-/-</sup> oocytes is suppressed by inactivation of SAC.** Oocytes from wild-type and *Meikin*<sup>-/-</sup> mice were cultured after GVBD in the presence of nocodazole (10  $\mu$ M) or the Mps1 kinase inhibitor reversin (5  $\mu$ M). The first polar body extrusion (PBE) rates are shown with mean  $\pm$  s.e.m. from 3 independent

experiments. The total number of oocytes is  $n = 27$  in wild-type with nocodazole (7, 10 and 10, respectively),  $n = 27$  for wild-type with reversine (6, 10 and 11, respectively),  $n = 21$  for *Meikin*<sup>-/-</sup> with nocodazole (5, 8 and 8, respectively) and  $n = 23$  for *Meikin*<sup>-/-</sup> with reversine (7, 8 and 8, respectively).

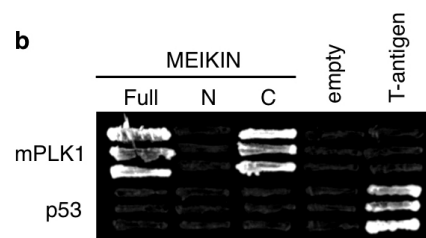


**Extended Data Figure 6 | MEIKIN is required for centromeric cohesion protection and mono-orientation in spermatocytes.** **a**, Squashed spermatocytes at metaphase I (Meta I) were immunostained for SGO2, ACA and DAPI. Partial z-axis projection images of aligned chromosomes are shown with magnified images of a bivalent (top). The signal intensity of SGO2 adjacent to the centromere was quantified and normalized to that of ACA. The relative intensities are shown with mean + s.e.m. from 3 independent experiments (bottom). In each experiment, 15 centromeres from a spermatocyte were quantified ( $n = 4$  cells). **b**, Squashed spermatocytes at anaphase I ( $5 \mu\text{m} < \text{segregated DNA mass distance} < 10 \mu\text{m}$ ) from wild-type, *Meikin*<sup>-/-</sup> and *Sgo2*<sup>-/-</sup> mice were immunostained for CENP-C, REC8 and DNA (top). A pair of sister kinetochores is magnified. The distance between sister kinetochores

was scored and represented in the scatter plot with median (bottom). A total of 15 kinetochores from 5 spermatocytes were measured in each group ( $n = \text{kinetochore number}$ ). **c**, Squashed spermatocytes at prometaphase I from wild-type and *Meikin*<sup>-/-</sup> mice, and MEF cells at prometaphase (prometa) were immunostained for CENP-C and DNA. Pairs of sister kinetochores are magnified. The distances between sister kinetochores were scored and represented in the graph with mean + s.e.m. from three independent experiments (right). Note that, MEF cells sample preparation and sister kinetochore distance measurements were performed same methods with spermatocytes. In each experiment, 10 kinetochores were measured in a cell ( $n = 5$  cells). \* $P < 0.05$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , unpaired  $t$ -test (a–c). Scale bars,  $5 \mu\text{m}$  (unless otherwise indicated).

**a**

Isolated clones	frequency	Isolated clones	frequency
<i>Rpl39l</i>	67	<i>Eif1</i>	1
<i>Dkk1</i>	30	<i>Epb4.1l3</i>	1
<i>BC089491</i>	21	<i>Gabarap</i>	1
<i>Kat5</i>	18	<i>1700023L04Rik</i>	1
<i>Morn2</i>	11	<i>eIF4GII</i>	1
<b>PLK1</b>	<b>10</b>	<i>Epb4.1l1</i>	1
<i>Ubb</i>	9	<i>Nsun7</i>	1
<i>Ggnbp1</i>	6	<i>Phf7</i>	1
<i>Ubl4b</i>	3	<i>4933400N17Rik</i>	1
<b>Cenpc1</b>	<b>2</b>	<i>Bace1</i>	1
<i>Mfap3l</i>	2	<i>EG433945</i>	1
<i>Ptpla</i>	2	<i>Snrpg</i>	1
<i>Stx8</i>	2	<i>Bxdc1</i>	1
<i>Dnajb8</i>	2	<i>Ints8</i>	1
<i>Tmem57</i>	2	<i>Ropn1</i>	1
<i>Gtsf1</i>	2	<i>Tctex1d1</i>	1
<i>Arhgap5</i>	1	<b>Total</b>	<b>207</b>

**b****c**

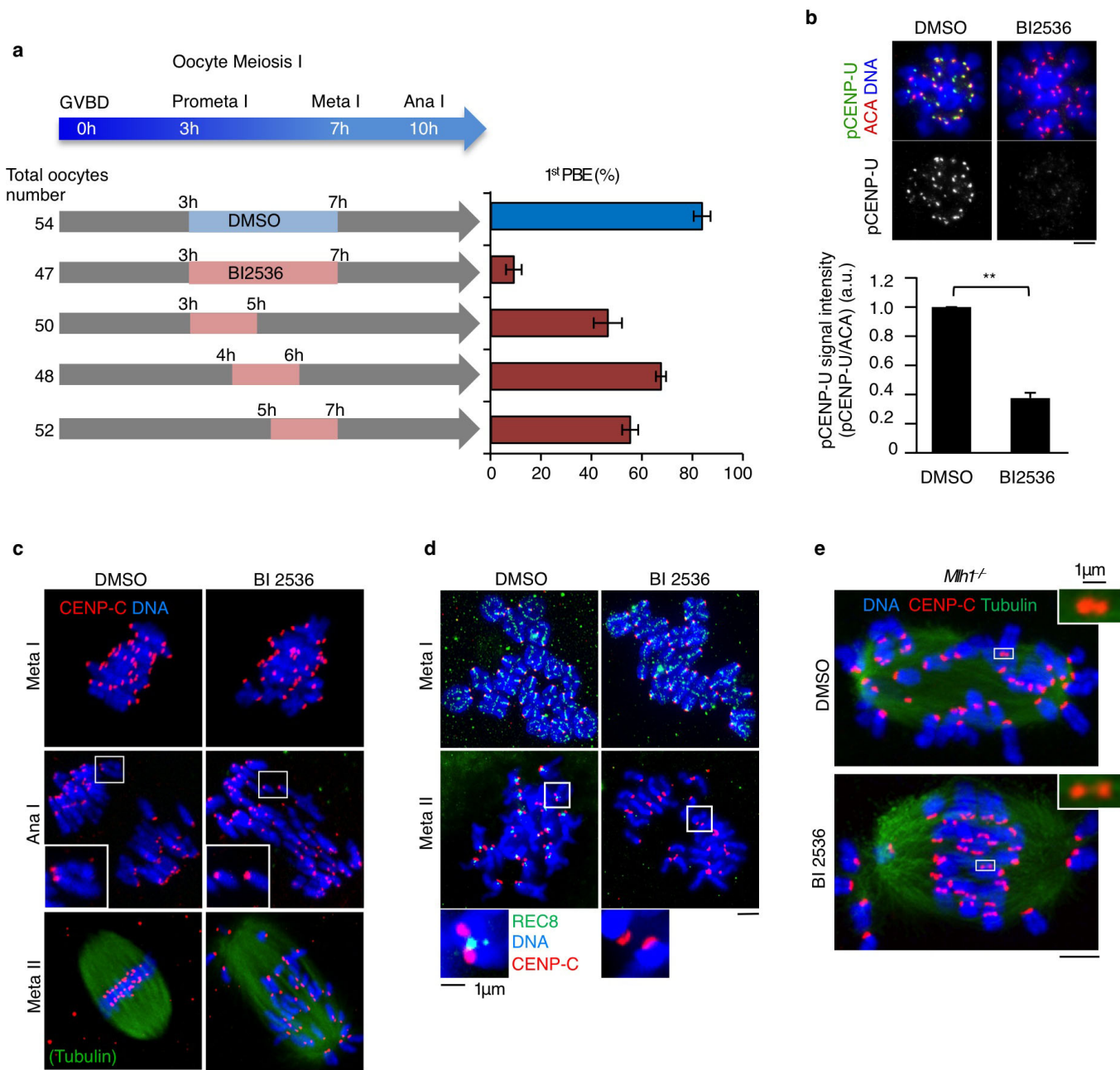
Reference	MW	Accession	A 1st	A 2nd	B 1st	B 2nd
			Peptide (Hits)	Peptide (Hits)	Peptide (Hits)	Peptide (Hits)
Scan(s)	Sp	RSp	lons	lons	lons	lons
thyroid hormone receptor associated protein 3	108113.5	68533246	3 (3 0 0 0 0)	2 (2 0 0 0 0)	11 (11 0 0 0 0)	12 (12 0 0 0 0)
splicing factor, arginine/serine-rich 10	33645.6	4759098	2 (2 0 0 0 0)	3 (3 0 0 0 0)	10 (10 0 0 0 0)	4 (4 0 0 0 0)
<b>4930404A10Rik (MEIKIN)</b>	<b>47393.8</b>	<b>58864863</b>	<b>14 (14 0 0 0 0)</b>	<b>6 (6 0 0 0 0)</b>	<b>10 (10 0 0 0 0)</b>	<b>9 (9 0 0 0 0)</b>
heterogeneous nuclear ribonucleoprotein M	77597.4	21313308	3 (3 0 0 0 0)	7 (7 0 0 0 0)	8 (8 0 0 0 0)	8 (8 0 0 0 0)
beta-tubulin	49608.0	12846758	7 (7 0 0 0 0)	10 (10 0 0 0 0)	5 (5 0 0 0 0)	8 (8 0 0 0 0)
matrin 3	94571.7	25141233	4 (4 0 0 0 0)	6 (6 0 0 0 0)	5 (5 0 0 0 0)	4 (4 0 0 0 0)
<b>Centromere protein C 1 (CENP-C)</b>	<b>102162.9</b>	<b>1345729</b>	<b>2 (2 0 0 0 0)</b>	<b>7 (7 0 0 0 0)</b>	<b>5 (5 0 0 0 0)</b>	<b>2 (2 0 0 0 0)</b>
mago-nashi homolog	17080.8	13385002	1 (1 0 0 0 0)	2 (2 0 0 0 0)	5 (5 0 0 0 0)	5 (5 0 0 0 0)
eukaryotic translation initiation factor 4A	46810.2	20149756	1 (1 0 0 0 0)	1 (1 0 0 0 0)	5 (5 0 0 0 0)	4 (4 0 0 0 0)
heterogeneous nuclear ribonucleoprotein F	45700.9	19527048	5 (5 0 0 0 0)	3 (3 0 0 0 0)	4 (4 0 0 0 0)	3 (3 0 0 0 0)
tubulin, alpha 8	50019.6	8394493	3 (3 0 0 0 0)	5 (5 0 0 0 0)	4 (4 0 0 0 0)	4 (4 0 0 0 0)
fragile X mental retardation-related protein 1	76175.3	163954941	2 (2 0 0 0 0)	3 (3 0 0 0 0)	4 (4 0 0 0 0)	6 (6 0 0 0 0)
nuclear receptor coactivator 5	65279.9	21450271	10 (10 0 0 0 0)	4 (4 0 0 0 0)	4 (4 0 0 0 0)	5 (5 0 0 0 0)
U5 snRNP-specific protein, 200 kDa	218284.4	50510635	1 (1 0 0 0 0)	2 (2 0 0 0 0)	4 (4 0 0 0 0)	5 (5 0 0 0 0)
DHX9_MOUSE ATP-dependent RNA helicase A	149379.8	71153505	3 (3 0 0 0 0)	3 (3 0 0 0 0)	3 (3 0 0 0 0)	3 (3 0 0 0 0)
tudor domain containing 1 isoform 4	102983.9	13878217	2 (2 0 0 0 0)	1 (1 0 0 0 0)	3 (3 0 0 0 0)	3 (3 0 0 0 0)
heterogeneous nuclear ribonucleoprotein H1	49168.4	10946928	1 (1 0 0 0 0)	1 (1 0 0 0 0)	3 (3 0 0 0 0)	4 (4 0 0 0 0)
mKIAA0324 protein [Mus musculus]	188565.5	28972153	1 (1 0 0 0 0)	1 (1 0 0 0 0)	3 (3 0 0 0 0)	4 (4 0 0 0 0)
PRP8	273424.6	67460967	1 (1 0 0 0 0)	2 (2 0 0 0 0)	3 (3 0 0 0 0)	3 (3 0 0 0 0)
RNA binding motif protein 14	69406.1	86262142	1 (1 0 0 0 0)	2 (2 0 0 0 0)	2 (2 0 0 0 0)	1 (1 0 0 0 0)
<b>polo-like kinase (PLK1)</b>	<b>68257.7</b>	<b>128485538</b>	<b>3 (3 0 0 0 0)</b>	<b>3 (3 0 0 0 0)</b>	<b>1 (1 0 0 0 0)</b>	<b>1 (1 0 0 0 0)</b>
YLP motif-containing protein 1 (YLP1)_	155033.5	18203404	2 (2 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)
unnamed protein product [Mus musculus]	57266.4	26338912	2 (2 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)
lactate dehydrogenase C	35889.1	7305229	1 (1 0 0 0 0)	2 (2 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)
interleukin enhancer binding factor 2	18673.7	26327701	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)
histone cluster 1, H2ah	13941.8	30061327	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)
heat shock protein 9	73415.7	162461907	1 (1 0 0 0 0)	4 (4 0 0 0 0)	1 (1 0 0 0 0)	1 (1 0 0 0 0)

**Extended Data Figure 7 | PLK1 was identified as a MEIKIN interactor.**

**a**, Yeast two-hybrid screening of mouse MEIKIN interactors was performed using a mouse testis cDNA library. The number of clones isolated by screening is summarized. Because the use of MEIKIN full-length and N-terminal region (amino acids 1–271) as bait resulted in a high background of false positive interactions in yeast two-hybrid screening, we used MEIKIN C terminus (amino acids 272–434) as bait. **b**, Yeast two-hybrid assay demonstrates that MEIKIN interacts directly with mouse PLK1 through the MEIKIN

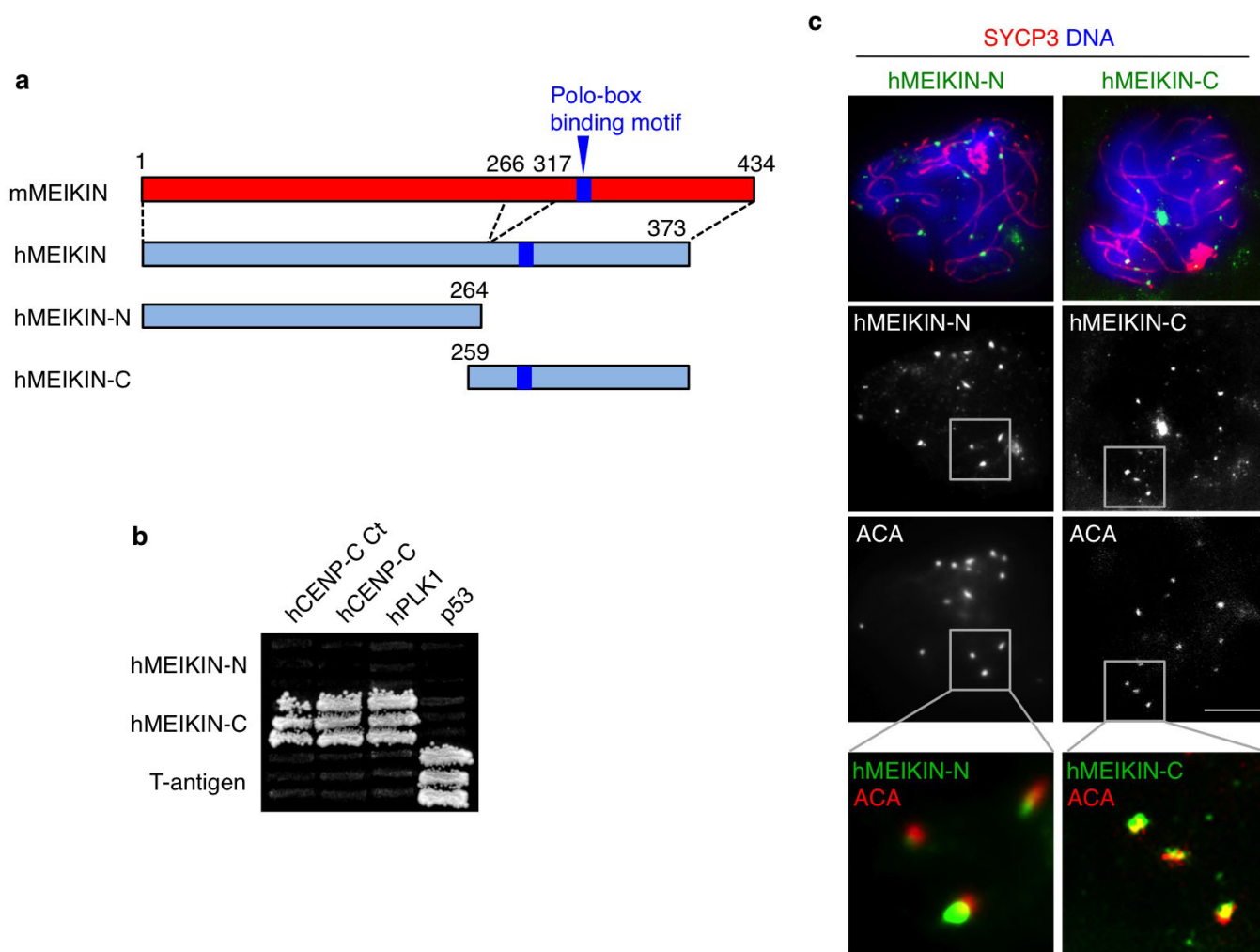
C-terminal domain (amino acids 272–434). **c**, Mouse MEIKIN protein was immunoprecipitated from testis chromatin-bound fraction by two different anti-MEIKIN polyclonal antibodies (A and B). The immunoprecipitates underwent two independent LC–MS/MS analyses. Those proteins commonly identified in all the LC–MS/MS analyses are listed with the number of peptide hits in the table. Note that polo-like kinase (PLK1) as well as CENP-C were repeatedly identified in the two-hybrid screening (**a**) and LC–MS/MS analyses (**c**).





**Extended Data Figure 8 | BI 2536 treatment reduces PLK1 kinase activity in oocytes.** **a**, Schematic illustration of BI 2536 treatment in wild-type oocyte culture (left). Oocytes were treated with DMSO or BI 2536 (100 nM) during the indicated time periods, then washed and released into normal culture medium. The first polar body extrusion ratio (1st PBE) was counted at 10 h after GVBD (right). Error bars, mean  $\pm$  s.e.m. from 3 independent experiments. The total number of oocytes used for each experiment is shown. **b**, Wild-type oocytes treated with DMSO or BI 2536 (during 6–7 h after GVBD) were fixed and immunostained for PLK1 substrate pCENP-U, ACA and DAPI at metaphase I (top). The relative pCENP-U intensity normalized to that of ACA is shown in the graph with mean  $\pm$  s.e.m. from 3 independent experiments. In each experiment, 10 kinetochores from an oocyte were quantified ( $n = 4$  cells). **c**, Oocytes were treated with DMSO or BI 2536

during the period of 4–6 h after GVBD and stained for  $\alpha$ -tubulin, CENP-C and DAPI (DNA) at the indicated stages in whole mount (related to Fig. 4d). Magnified images are shown to highlight the separation of sister kinetochores in BI 2536-treated oocyte in anaphase I. **d**, Chromosome spreads from control and BI 2536-treated wild-type oocytes at metaphase I (6 h after GVBD) and metaphase II (20 h after GVBD) were stained for REC8, CENP-C and DAPI (DNA) (related to Fig. 4d). Magnified images are shown to highlight the loss of cohesion in BI 2536 treated oocytes in metaphase II. **e**, *Mlh1*<sup>-/-</sup> oocytes (C57BL/6 background) were treated with BI 2536 between 9–10 h after GVBD and stained for  $\alpha$ -tubulin, CENP-C and DAPI (DNA) in whole mount (related to Fig. 4e). Magnified images are shown to highlight bi-oriented sister kinetochores at the spindle midzone. Scale bars, 5  $\mu$ m (unless otherwise indicated).



**Extended Data Figure 9 | The human homologue of MEIKIN.** **a**, Schematic illustrations of mouse and human MEIKINs. The putative amino acid sequence of hMEIKIN full-length (373 amino acids) was deduced from our own sequencing of DNA (see Extended Data Fig. 2), which was amplified by RT-PCR from the human testis cDNA. The amino acid sequence equivalent to mouse exon 10 (mEX10) is absent in the hMEIKIN protein, despite our attempts at a computational search to identify the missing DNA sequence. **b**, Yeast two-hybrid assays demonstrate that the hMEIKIN C terminus interacts

with human CENP-C (full length), hCENP-C C-terminal (CENPC motif + Mif2 motif, amino acids 732–945) and hPLK1, in agreement with mouse MEIKIN data. **c**, Immunostaining of a human seminiferous tubule section (purchased from Biochain) demonstrates that hMEIKIN localizes to centromeres (ACA) in pachytene spermatocytes. We used anti-hMEIKIN-N and anti-hMEIKIN-C antibodies. Enlarged images of the rectangles are shown to highlight the co-localization of ACA and hMEIKIN (bottom). Scale bar, 5  $\mu$ m.

# Long-lived magnetism from solidification-driven convection on the pallasite parent body

James F. J. Bryson<sup>1</sup>, Claire I. O. Nichols<sup>1</sup>, Julia Herrero-Albillos<sup>2,3</sup>, Florian Kronast<sup>4</sup>, Takeshi Kasama<sup>5</sup>, Hossein Alimadadi<sup>5</sup>, Gerrit van der Laan<sup>6</sup>, Francis Nimmo<sup>7</sup> & Richard J. Harrison<sup>1</sup>

**Palaeomagnetic measurements of meteorites<sup>1–5</sup> suggest that, shortly after the birth of the Solar System, the molten metallic cores of many small planetary bodies convected vigorously and were capable of generating magnetic fields<sup>6</sup>. Convection on these bodies is currently thought to have been thermally driven<sup>7,8</sup>, implying that magnetic activity would have been short-lived<sup>9</sup>. Here we report a time-series palaeomagnetic record derived from nanomagnetic imaging<sup>10</sup> of the Imilac and Esquel pallasite meteorites, a group of meteorites consisting of centimetre-sized metallic and silicate phases. We find a history of long-lived magnetic activity on the pallasite parent body, capturing the decay and eventual shutdown of the magnetic field as core solidification completed. We demonstrate that magnetic activity driven by progressive solidification of an inner core<sup>11–13</sup> is consistent with our measured magnetic field characteristics and cooling rates<sup>14</sup>. Solidification-driven convection was probably common among small body cores<sup>15</sup>, and, in contrast to thermally driven convection, will have led to a relatively late (hundreds of millions of years after accretion), long-lasting, intense and widespread epoch of magnetic activity among these bodies in the early Solar System.**

The pallasites are slowly cooled (2–9 K per million years, Myr; ref. 14) meteorites consisting of olivine crystals embedded in an Fe–Ni matrix<sup>16</sup>, which have been suggested to originate from the mid to upper mantle of a ~200-km-radius body<sup>1</sup>. These slow cooling rates allowed for characteristic microstructures to form in the metallic matrices<sup>17</sup>, such as regions of intergrown nanometre-scale islands of tetraenaite (ordered FeNi)<sup>18,19</sup> and an ordered Fe<sub>3</sub>Ni matrix<sup>20</sup>, collectively known as cloudy zones (CZ)<sup>21</sup>. During the cooling of the parent body, these tetraenaite islands exsolved and subsequently coarsened over tens of millions of years<sup>22</sup>. The island diameter decreases systematically across the CZ, reflecting a decrease in the local formation age of the islands<sup>20</sup>. Each island has adopted one of three orthogonal magnetic easy axes as it formed<sup>18,23</sup>, and thus could display any one of six magnetization directions. Variations in the intensity and direction of an external magnetic field have led to measurable differences in the populations of each magnetization direction<sup>20</sup>. Crucially, the temporal evolution of an external field is recorded by the variations in the relative proportions of these directions across the CZ<sup>10</sup>, which can be quantified using high-resolution nanomagnetic imaging. Such images were captured for the Imilac and Esquel pallasites, using X-ray magnetic circular dichroism<sup>24,25</sup> at the X-ray photoemission electron microscope (XPEEM)<sup>26</sup> at the BESSY II synchrotron in Berlin, which provides the spatially resolved magnetization of a sample surface with a resolution down to 40 nm over a 5  $\mu$ m field-of-view<sup>10</sup>.

Four and six non-overlapping, 450-nm-wide regions across the CZ (decreasing age) were extracted from the XPEEM images of the Imilac and Esquel meteorites, respectively (Fig. 1). The field recorded by each region was deduced by comparing the experimental XPEEM signal to that of simulated CZ nanostructures magnetized by variable field components<sup>10</sup> (see Methods).

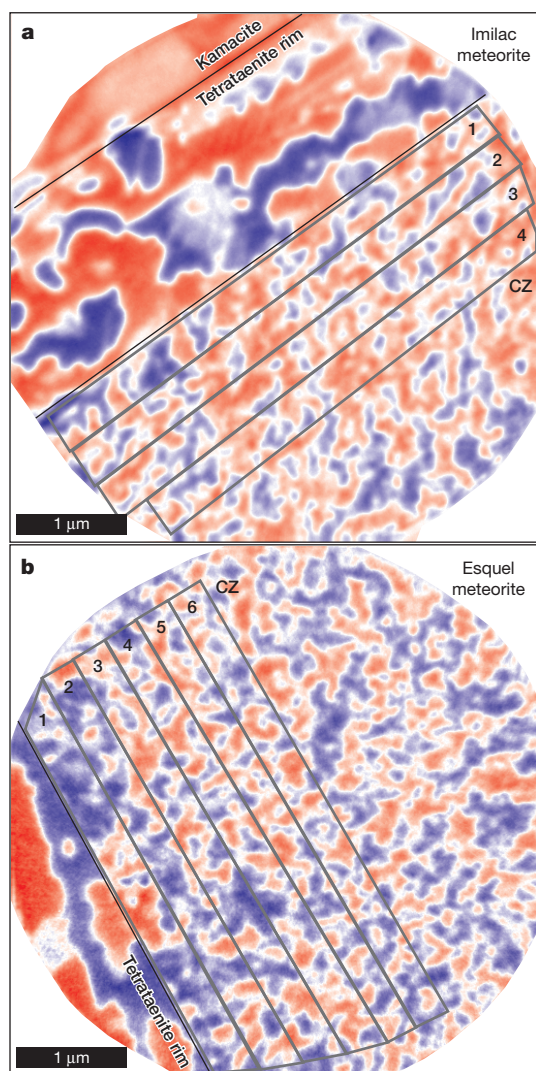
The field components that created the five best fits to the experimental data for all regions in the Imilac meteorite and regions 1–5 in the Esquel meteorite are self-consistent, but those of region 6 in the Esquel meteorite display a large degree of scatter (Fig. 2). These observations imply that the time-averaged field experienced by each region is unidirectional. The Imilac meteorite recorded a roughly constant intensity between  $119(\pm 12)$   $\mu$ T and  $131(\pm 13)$   $\mu$ T across all regions (Fig. 3a). The Esquel meteorite recorded a very different trend, with an initial intensity of  $84 \pm 14$   $\mu$ T, which decreases down to a plateau at a value of  $31(\pm 10)$ – $35(\pm 7)$   $\mu$ T. Owing to the limitations of the fitting, the field recorded by region 6 could be constrained only to  $\leq 10$   $\mu$ T (Fig. 3b), which, coupled with the large degree of scatter, suggests that this region may not have experienced a field. These intensity values are consistent with those reported independently<sup>1</sup>, deduced by a different method. The main uncertainty in our intensity values is the island volume when the field was recorded (see Supplementary Information). With the largest possible (present-day) island volume, the intensities of the first regions of the Imilac and Esquel meteorites are  $4 \pm 0.4$   $\mu$ T and  $2 \pm 0.7$   $\mu$ T, respectively. Our main result, however, is the inferred time-variation in relative field strength and orientation, which is unaffected by the absolute recording volume.

To estimate when each meteorite recorded the magnetic field, we modelled the cooling of a 200-km-radius<sup>1</sup> conductive body during the ~250 Myr following accretion (Fig. 4). Depths were assigned to the Imilac and Esquel meteorites, based on their cooling rates at 800 K (Extended Data Fig. 6), calculated using the observed island sizes of  $147 \pm 4$  nm and  $158 \pm 6$  nm, respectively (Extended Data Fig. 2), and the empirical relationship between island size and cooling rate<sup>14</sup>. The Imilac and Esquel meteorites resided at depths of  $38 \pm 2.5$  km and  $45 \pm 4$  km, respectively. The CZ started recording the field once tetraenaite formation commenced (593 K), which we propose occurred at these depths during the period of core solidification (Fig. 4).

At this time, dynamo activity could have been driven by compositional convection resulting from the preferential fractionation of light elements (for example, S) into the liquid outer core as the inner core solidified<sup>11</sup>, as is the case for the present-day geodynamo<sup>6,12</sup>. To assess this hypothesis, the magnetic moment, magnetic Reynolds number and local Rossby number (Extended Data Fig. 7) were calculated<sup>13</sup> for a compositionally driven dynamo throughout the period of core solidification (Fig. 3c). In our model (see Supplementary Information), the thermal structure of the mantle governs the core–mantle boundary heat flux, which in turn dictates the rates of inner-core growth and light-element rejection that ultimately control convection<sup>15</sup>. For the majority of core solidification, the core magnetic moment is predicted to decrease smoothly to zero. The magnetic Reynolds number governs whether the convective motion generates a field, and, for a four-hour rotation period, dynamo activity is predicted to have ceased at the latest stages of core solidification. The local Rossby number governs the

<sup>1</sup>Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK. <sup>2</sup>Centro Universitario de la Defensa, Carretera de Huesca s/n, E-50090 Zaragoza, Spain. <sup>3</sup>Instituto de Ciencia de Materiales de Aragón, CSIC—Universidad de Zaragoza, Pedro Cerbuna 12, E-50009 Zaragoza, Spain. <sup>4</sup>Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, 12489 Berlin, Germany. <sup>5</sup>Center for Electron Nanoscopy, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark. <sup>6</sup>Diamond Light Source, Chilton, Didcot, Oxfordshire OX11 0DE, UK. <sup>7</sup>Department of Earth and Planetary Sciences, University of California, Santa Cruz, California 95064, USA.

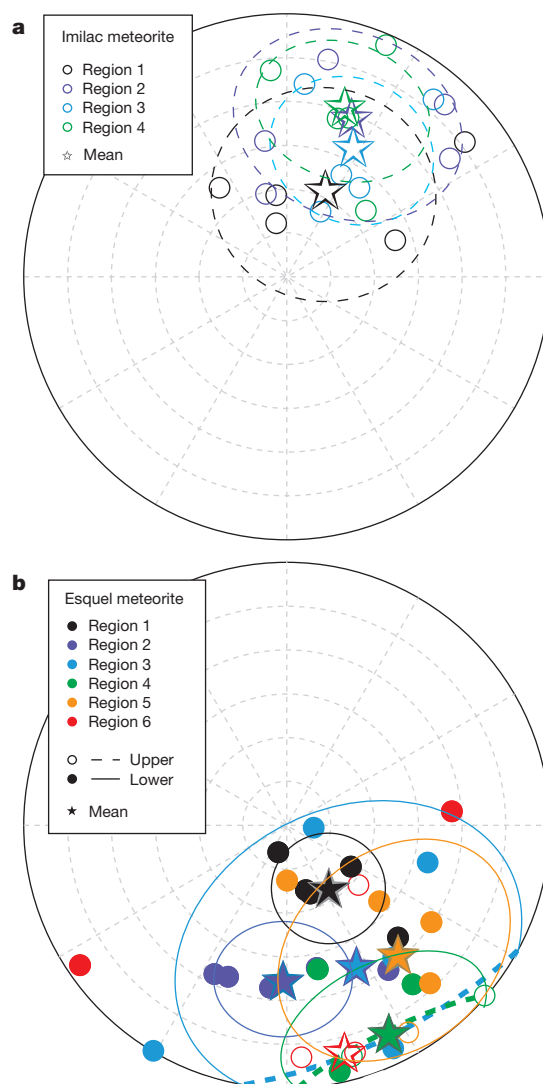




**Figure 1 | Representative XPEEM images of the kamacite, tetrataenite rim and CZ in the Imilac and Esquel pallasites.** Blue and red signals correspond to positive and negative projections of the magnetization along the X-ray beam direction (Extended Data Fig. 1) in the Imilac (a) and Esquel (b) meteorites. Each image is one of four used in the palaeomagnetic analysis (Extended Data Figs 3 and 4). The CZ displays a complex interlocking pattern of positive and negative domains. The CZ regions are labelled. The age of the CZ decreases with distance from the tetrataenite rim.

polarity of the moment, and, for the same rotation period, we predict a dipolar to multipolar transition slightly before dynamo cessation. This transition is accompanied by a decrease of up to a factor of 20 in the field intensity<sup>13</sup>.

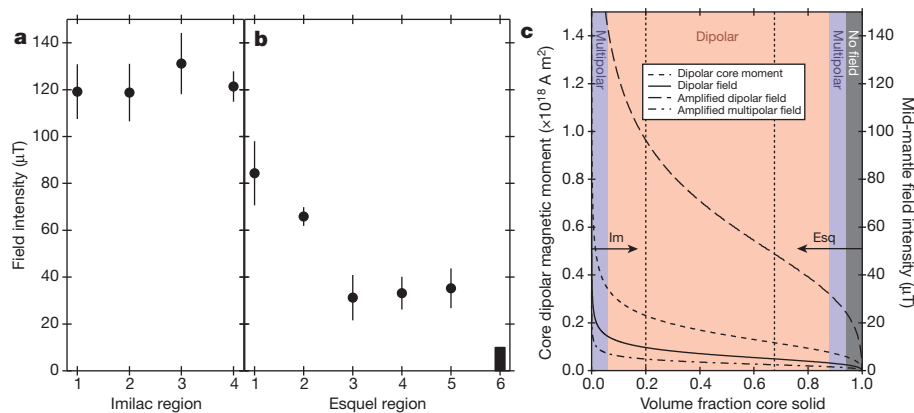
The conductive cooling model predicts that the Imilac and Esquel meteorites started recording the dynamo field during the early (sometime between when 0% and 33% of the core volume had solidified) and late (when 43% to 100% of the core volume had solidified) stages of inner-core growth, respectively (Fig. 4). In each case, recording is estimated to have continued while a further 10%–20% of core volume solidified. The Imilac meteorite is therefore predicted to have experienced a smoothly decreasing dipolar dynamo field (corresponding to the early decreasing core moment), whereas the Esquel meteorite is predicted to have captured a decrease in the field intensity (dipolar–multipolar transition), a constant weak intensity (multipolar regime), and finally zero intensity (dynamo cessation). Both trends agree, within uncertainty, with the results inferred from their respective meteorites (Fig. 3), indicating that the recorded fields were generated by compositional convection.



**Figure 2 | The five best-fitting and average field components for each CZ region of the Imilac and Esquel pallasites.** Equal-area stereographic projection of the field components in the Imilac (a) and Esquel (b) meteorites. Empty points and dashed lines represent the lower hemisphere of the stereoplot, and filled circles and solid lines represent the upper hemisphere. The 95% confidence interval, assuming a Fisher distribution, is included as an ellipse around each of the average directions. This ellipse is not included for region 6 of the Esquel meteorite owing to the large degree of scatter.

The model dynamo field intensity midway through the mantle (corresponding roughly to the depths of the two meteorites) (Fig. 3c) is about an order of magnitude less than the nominal values extracted from the CZ. This discrepancy may be explained through the amplification of the dynamo field by the mantle-hosted metal that constitutes the pallasites. At the CZ magnetization acquisition temperatures, this metal will have consisted predominantly of the magnetically soft phases kamacite and taenite<sup>27</sup>. By considering the extremes in feasible metal morphologies (see Supplementary Information), the field within the metal will be amplified by a factor of 3–200 (Fig. 3c), generating intensities well within the range of our measured values.

The values of many of the parameters in the cooling and dynamo scaling models are uncertain (for example, the light-element content of the core liquid and variation in this parameter during solidification). Also, the scaling model is derived empirically from numerical models with parameters different to those relevant to the interior of small bodies,



**Figure 3 | Measured and simulated dynamo field intensity trends.**

Measured intensities for the Imilac (a) and Esquel (b) meteorites. Points and error bars are the means and standard deviations, respectively, of the five best-fitting field components. The  $\leq 10 \mu\text{T}$  field for region 6 of the Esquel meteorite is depicted as a black rectangle. Each region corresponds to a

maximum time period of  $\sim 1\text{--}2$  Myr. c, Simulated mid-mantle intensities.

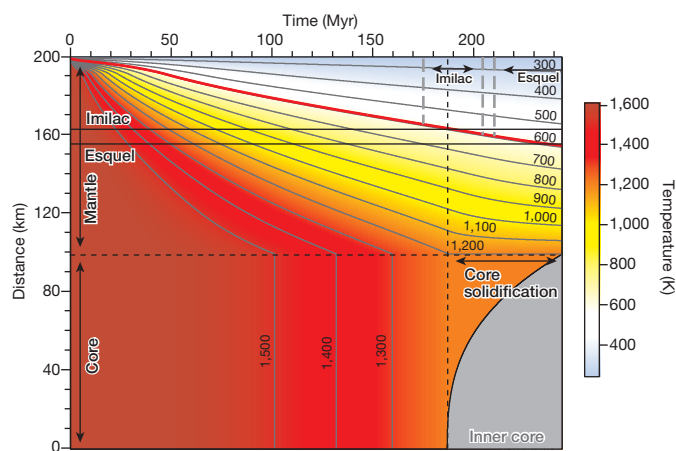
Dipolar amplification:  $10 \times$  dipolar field. Multipolar amplification:  $0.05 \times$  amplified dipolar field. 'Im' and 'Esq' mark the recording periods of the Imilac and Esquel meteorites, respectively (from Fig. 4).

and not all small-body cores are expected to have solidified from the bottom up<sup>28</sup>. Despite these caveats, the field intensity trends inferred from the CZ of the pallasites agree with our model predictions. The thermally driven dynamo activity that is widely believed to have acted on small bodies<sup>7–9</sup> is inefficient<sup>15</sup> (and was therefore probably uncommon) and could only have generated fields for the first 10–50 Myr (depending on the body's radius) after the formation of the Ca–Al-rich inclusions (CAIs)<sup>7,29</sup>. Our results suggest that compositionally driven dynamo activity was generated in outwardly solidifying small-body cores from  $\sim 60$  Myr to 250 Myr after CAI formation, and lasted for the majority of core solidification (a further  $\sim 25\text{--}150$  Myr, depending on the body's radius)<sup>1</sup>. These later fields were extremely easy to generate<sup>15</sup> (and so would have been widespread), relatively constant in both direction and intensity, and were capable of displaying both dipolar and multipolar morphologies. These conclusions imply that a second epoch of dynamo activity across a potentially large fraction of small bodies occurred in the early Solar System, and help to explain the long-lived magnetic activity observed for other bodies, such as the Moon<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 8 September; accepted 21 November 2014.**

1. Tarduno, J. A. *et al.* Evidence for a dynamo in the main group pallasite parent body. *Science* **338**, 939–942 (2012).
2. Carporzen, L. *et al.* Magnetic evidence for a partially differentiated carbonaceous chondrite parent body. *Proc. Natl Acad. Sci. USA* **108**, 6386–6389 (2011).
3. Fu, R. R. *et al.* An ancient core dynamo in asteroid Vesta. *Science* **338**, 238–241 (2012).
4. Weiss, B. P., Fong, L. E., Vali, H., Lima, E. A. & Baudenbacher, F. J. Paleointensity of the ancient Martian magnetic field. *Geophys. Res. Lett.* **35**, L23207 (2008).
5. Garrick-Bethell, I., Weiss, B. P., Shuster, D. L. & Buz, J. Early lunar magnetism. *Science* **323**, 356–359 (2009).
6. Stevenson, D. J. Planetary magnetic fields: achievements and prospects. *Space Sci. Rev.* **152**, 651–664 (2010).
7. Elkins-Tanton, L. T., Weiss, B. P. & Zuber, M. T. Chondrites as samples of differentiated planetesimals. *Earth Planet. Sci. Lett.* **305**, 1–10 (2011).
8. Weiss, B. P. *et al.* Magnetism on the angrite parent body and the early differentiation of planetesimals. *Science* **322**, 713–716 (2008).
9. Weiss, B. P., Gattacceca, J., Stanley, S., Rochette, P. & Christensen, U. R. Paleomagnetic records of meteorites and early planetesimal differentiation. *Space Sci. Rev.* **152**, 341–390 (2010).
10. Bryson, J. F. J. *et al.* Nanopaleomagnetism of meteoritic Fe–Ni studied using X-ray photoemission electron microscopy. *Earth Planet. Sci. Lett.* **396**, 125–133 (2014).
11. Fearn, D. R. & Loper, D. E. Compositional convection and stratification of Earth's core. *Nature* **289**, 393–394 (1981).
12. Nimmo, F. in *Treatise on Geophysics* (ed. Schubert, G.) 8.02, 31–65 (Elsevier, 2007).
13. Olson, P. & Christensen, U. R. Dipole moment scaling for convection-driven planetary dynamos. *Earth Planet. Sci. Lett.* **250**, 561–571 (2006).
14. Yang, J., Goldstein, J. I. & Scott, E. R. D. Main-group pallasites: thermal history, relationship to IIIAB irons, and origin. *Geochim. Cosmochim. Acta* **74**, 4471–4492 (2010).
15. Nimmo, F. Energetics of asteroid dynamos and the role of compositional convection. *Geophys. Res. Lett.* **36**, L10210 (2009).
16. Anders, E. Origin, age and composition of meteorites. *Space Sci. Rev.* **3**, 583–714 (1964).
17. Goldstein, J. I., Scott, E. R. D. & Chabot, N. L. Iron meteorites: crystallization, thermal history, parent bodies, and origin. *Chem. Erde Geochem.* **69**, 293–325 (2009).
18. Néel, L., Pauleve, J., Pauthenet, R., Laugier, J. & Dautreppe, D. Magnetic properties of nickel-iron alloys bombarded by neutrons in a magnetic field. *J. Appl. Phys.* **35**, 873–876 (1964).
19. Clarke, R. S. & Scott, E. R. D. Tetraenaite-ordered FeNi, a new mineral in meteorites. *Am. Mineral.* **65**, 624–630 (1980).
20. Bryson, J. F. J., Church, N. S., Kasama, T. & Harrison, R. J. Nanomagnetic intergrowths in Fe–Ni meteoritic metal: the potential for time-resolved records of planetesimal dynamo fields. *Earth Planet. Sci. Lett.* **388**, 237–248 (2014).
21. Scott, E. R. D. The nature of dark-etching rims in meteoritic taenite. *Geochim. Cosmochim. Acta* **37**, 2283–2294 (1973).
22. Uehara, M., Gattacceca, J., Leroux, H., Jacob, D. & van der Beek, C. J. Magnetic microstructures of metal grains in equilibrated ordinary chondrites and implications for paleomagnetism of meteorites. *Earth Planet. Sci. Lett.* **306**, 241–252 (2011).
23. Albertsen, J. F. Tetragonal lattice of tetraenaite (ordered Fe–Ni, 50–50) from 4 meteorites. *Phys. Scr.* **23**, 301–306 (1981).



**Figure 4 | Cooling of a 200-km-radius body, consisting of a core, mantle and regolith.** Evolution of the temperature (colour and contours) at a given radius and time. The red contour depicts the initial tetraenaite formation temperature (593 K), corresponding to the earliest CZ remanence. Black horizontal lines mark the inferred depths of the Imilac and Esquel meteorites, calculated<sup>14</sup> from the average largest tetraenaite island size (Extended Data Fig. 2). Dashed grey vertical lines mark the likely recording period for each meteorite, inferred from the size range of largest CZ islands. During solidification (grey region), the core was isothermal at 1,200 K.

24. Stöhr, J. Exploring the microscopic origin of magnetic anisotropies with X-ray magnetic circular dichroism (XMCD) spectroscopy. *J. Magn. Magn. Mater.* **200**, 470–497 (1999).
25. van der Laan, G. Applications of soft X-ray magnetic dichroism. *J. Phys. Conf. Ser.* **430**, 012127 (2013).
26. Stöhr, J., Padmore, H., Anders, S., Stämmler, T. & Scheinfein, M. Principles of X-ray magnetic dichroism spectromicroscopy. *Surf. Rev. Lett.* **5**, 1297–1308 (1998).
27. Yang, J. & Goldstein, J. I. The formation of the Widmanstätten structure in meteorites. *Meteorit. Planet. Sci.* **40**, 239–253 (2005).
28. Williams, Q. Bottom-up versus top-down solidification of the cores of small solar system bodies: constraints on paradoxical cores. *Earth Planet. Sci. Lett.* **284**, 564–569 (2009).
29. Sterenborg, M. G. & Crowley, J. W. Thermal evolution of early solar system planetesimals and the possibility of sustained dynamos. *Earth Planet. Sci. Lett.* **214**, 53–73 (2013).
30. Laneuville, M. *et al.* A long-lived lunar dynamo powered by core crystallization. *Earth Planet. Sci. Lett.* **401**, 251–260 (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We acknowledge the Helmholtz-Zentrum Berlin for the use of the synchrotron radiation beamtime at beamline UE49 of BESSY II. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC grant agreement numbers 320750 and 312284, the Natural Environment Research Council, Fundación ARAID and the Spanish MINECO MAT2011-23791. We thank the Natural History Museum, London, and the Sedgwick Museum of Earth Sciences, University of Cambridge, for samples. We also thank J. Tarduno for discussions concerning magnetic shielding.

**Author Contributions** J.F.J.B., C.I.O.N., J.H.-A., F.K., G.v.d.L. and R.J.H. collected the XPEEM images. T.K. and H.A. collected the SEM images. J.F.J.B. and C.I.O.N. analysed the XPEEM images. J.F.J.B. and F.N. performed the planetary cooling and dynamo generation simulations. J.F.J.B., R.J.H. and F.N. wrote the paper with contributions from all other authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.F.J.B. ([jfjb2@cam.ac.uk](mailto:jfjb2@cam.ac.uk)).



## METHODS

**Sample characterization and experimental details.** The Imilac meteorite sample was provided by the Sedgwick Museum of Earth Sciences, University of Cambridge, sample number 11587. The Esquel meteorite sample was provided by the Natural History Museum, London, sample number BM.1964.65. Both samples were cut and polished to generate the flat, clean surface required for XPEEM measurements. Both samples contained multiple, large (centimetre-sized) olivine crystals, which were surrounded by swathing kamacite. The CZ bordering swathing kamacite was studied in both meteorites. Care was taken at all stages that the experimental team were in possession of the samples to limit the exposure of the samples to both natural and artificial magnetic fields.

The XPEEM measurements were performed at the SPEEM UE49 end station at the BESSY II synchrotron. The sample surfaces were illuminated with a monochromatic, circularly polarized, X-ray beam tuned to the Fe  $L_{2,3}$  edge. A 10-kV or 15-kV high voltage was used to produce images with a high spatial resolution, which varied between 100 nm and 200 nm. We believe that this relatively poor resolution (XPEEM can achieve at best 30–40 nm resolution) is sample-specific and due to strong stray fields emanating from the sample surface as well as surface roughness. The XPEEM signal is enhanced by capturing and subtracting images with opposite circular polarizations<sup>10</sup>. Secondary electrons can only escape from the top ~5 nm of the sample<sup>31</sup>, so XPEEM is only capable of probing the surface magnetization of the sample. The beam was orientated 16° out of the plane of the sample, and 16° clockwise from the vertical edge of the image (Extended Data Fig. 1). The analysed images have a 5  $\mu$ m field of view and a resolution of ~10 nm per pixel. During polishing, the top 80–100 nm of the sample was physically altered such that the magnetization was unrepresentative of the underlying natural signal<sup>10,20</sup>. To remove this layer, the meteorites were argon ion sputtered under ultrahigh vacuum: the Esquel meteorite was sputtered for 12 h at 0.8 keV, followed by 3 h at 0.4 keV and the Imilac meteorite was sputtered for 16.5 h at 0.8 keV followed by 1 h at 0.4 keV).

Scanning electron microscopy (SEM) backscattered electron images were captured at the Center for Electron Nanoscopy, Technical University of Denmark, on a FEI Helios NanoLab 600. The images were acquired at an electron probe current of 5.5 nA and an acceleration voltage of 3 kV. SEM was used to measure the diameter of the tetraenite islands across the CZ in both meteorites (Extended Data Fig. 2). The island diameter was used to calculate the size of the blurring function during palaeomagnetic analysis. The average island size for each region is presented in Extended Data Table 1.

**Palaeomagnetic analysis.** The 5  $\mu$ m field-of-view XPEEM images were captured at multiple locations across both samples. The quality and orientation of each image was assessed and only those of high quality (that is, showing no beam drift, oxidation or charging) and those with features at a constant orientation to the X-ray beam were further analysed. These selection criteria resulted in four images for each meteorite. Pixel-intensity histograms were extracted from adjacent, non-overlapping, 450-nm-wide regions of all four XPEEM images of each meteorite. Histograms of the same region from all four images of the same meteorite were averaged to generate a representative data set for each region (Extended Data Figs 3 and 4).

The strength and direction of the field recorded by each region were deduced by comparing the measured XPEEM signal to that of a simulated nanostructure magnetized by variable field components<sup>10</sup>. First, the XPEEM intensities corresponding to each of the six possible magnetization directions were extracted from the tetraenite rim (positive and negative values of the three distinct twin domain intensities). These intensities were used to populate the islands in a simulated CZ nanostructure. This nanostructure consisted of 800 islands, created by eroding Voronoi cells<sup>10</sup>. The matrix is modelled as a soft magnet<sup>20</sup>, and its XPEEM signal is calculated assuming strong exchange coupling to the tetraenite islands<sup>10</sup>. The combined island and matrix signal was then convolved with an approximation to the experimental point spread function (calculated from the width of the domain boundaries in the tetraenite rim and the measured CZ island size, Extended Data Fig. 2), and noise was added.

To identify the properties of the field recorded by each region, the XPEEM signals expected for a range of field intensities and directions were simulated and compared to the experimental data, according to the procedure outlined in Extended Data Fig. 5. First, an approximate field intensity was identified for each region using a trial-and-error method, where the field components were altered manually until the simulated pixel-intensity histogram was at least in partial agreement with the average experimental histogram. A direction was then chosen at random from a Gaussian distribution. This direction was scaled to the approximate intensity value, and the proportions of the magnetization directions expected for this field were calculated<sup>20</sup> for a 45-nm island radius (see Supplementary Information). The simulated nanostructure was populated with these proportions and the corresponding pixel-intensity histogram calculated. This process was repeated three times, and the individual histograms were used to create an average histogram representative of the field components. The squared difference  $\chi^2$  between the average simulated histogram

and the experimental histogram was then calculated. This whole procedure was repeated to generate  $\chi^2$  values corresponding to 150 randomly selected directions. We identified the five directions with the lowest  $\chi^2$  values. These five directions correspond roughly to that of the field experienced by each region. These five field components were then used as the inputs for full least-squares fits to the experimental pixel-intensity histogram for each region. In this fitting procedure, the values of all three components were varied systematically (allowing both the field intensity and direction to vary) to create simulated nanostructures that minimized the  $\chi^2$  value between the simulated and experimental pixel intensity histograms. Again, the simulated nanostructures were populated three times for each set of field components. This process resulted in five sets of field components that all produced very high-quality fits to the experimental data (Fig. 2).

**Planetary cooling model.** The planetary cooling model employs multiple simplifications and there are large uncertainties in many of the values of the variable used; hence, we adopt a very simple approach<sup>1</sup>. The equations are all included in the Supplementary Information.

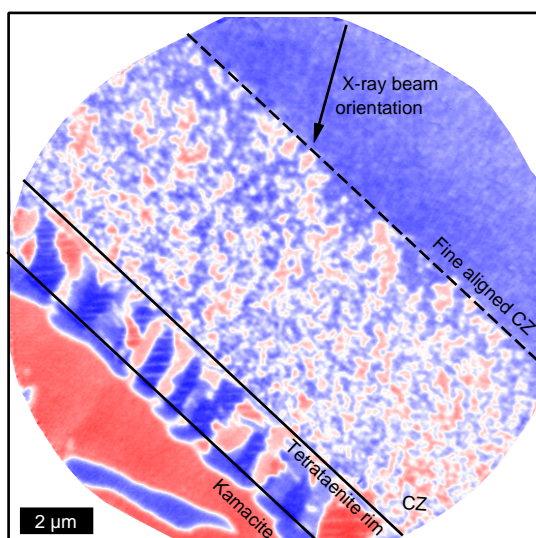
Initially, the entire body is isothermal at 1,600 K, which is the approximate silicate solidus. The parent body is assumed to cool rapidly to this temperature via advection of melt to the surface. Once this advective phase has ended, subsequent cooling will be much slower because it is controlled by conduction rather than melt advection. We further assume that no extra sources of heat were present (such as long-lived nuclides or heating from impacts). The temperature at the surface is kept at a constant value of  $T_s = 250$  K (ref. 32). The radius of the body is taken as 200 km (ref. 1), which places the pallasites in the mid to upper mantle on the basis of their measured cooling rates at 800 K (ref. 10). We also assumed an 8-km-thick megaregolith at the asteroid surface with a thermal diffusivity an order of magnitude lower than that of the mantle<sup>33,34</sup>. The upper half of the asteroid radius was modelled as the mantle, and the lower half as the core. Our cooling model captures the periods before and during core solidification. The core liquid is assumed to convect, so at any given time step is modelled as isothermal. Before solidification, the core temperature was modified from the value at the previous time step on the basis of the heat extracted across the core–mantle boundary during that time step. The core started to solidify when it first reached 1,200 K (ref. 1). Beyond this time, the core temperature was manually held at 1,200 K to approximate the fact that the core temperature barely changes during solidification, owing to the small slope of the melting curve. Core solidification was modelled as complete once the total heat extracted across the core–mantle boundary exceeded the total amount produced by core solidification. The values of the parameters used in the model are presented in Extended Data Table 2.

The cooling rate was found by differentiating the cooling model results with respect to time (Extended Data Fig. 6). Original depths in the parent body were determined for the pallasites by comparing their inferred cooling rates at 800 K with those calculated from Fig. 4. The experimental cooling rates were calculated using island sizes of  $147 \pm 4$  nm and  $158 \pm 6$  nm for the Imilac and Esquel meteorites, respectively (Extended Data Fig. 2), and the relationship between island size and cooling rate<sup>14</sup>. Cooling models were also performed for 100-km-radius and 300-km-radius parent bodies, to confirm that the value of 200 km used in this study is realistic (Extended Data Fig. 8).

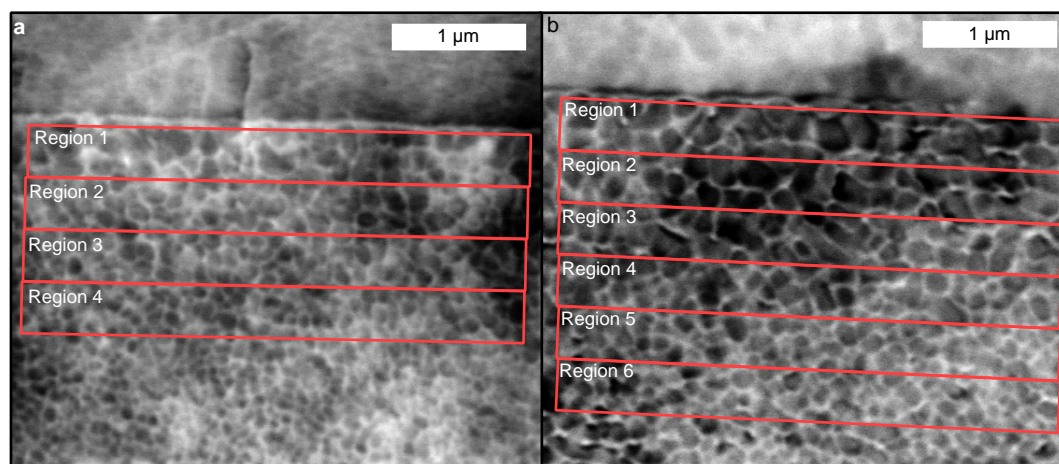
**Dynamo generation model.** To investigate the possibility of thermally driven dynamo activity, the cooling rate of the core liquid during solidification was calculated<sup>15</sup> from the heat flux across the core–mantle boundary (using temperature values from Fig. 4). All details and equations are outlined in the Supplementary Information.

To investigate compositional convection, the rate of inner core growth was calculated from the liquid cooling rate. This was performed iteratively with a time step of  $2 \times 10^{10}$  s from an initial core radius of 5 km up to the entire radius of the core (100 km). The growth rate was in turn used to calculate the buoyancy flux generated by the ejection of light elements during solidification (using a fluid density contrast of  $195 \text{ kg m}^{-3}$ ). This parameter was then used to calculate the flux-based Rayleigh number<sup>13</sup>. From this parameter, the local Rossby number, magnetic Reynolds number and dipolar Lorentz number were all calculated (Extended Data Fig. 7). All equations are presented in the Supplementary Information.

- Ohldag, H. *et al.* Spectroscopic identification and direct imaging of interfacial magnetic spins. *Phys. Rev. Lett.* **87**, 247201 (2001).
- Hevey, P. J. & Sanders, I. S. A model for planetesimal meltdown by  $^{26}\text{Al}$  and its implications for meteorite parent bodies. *Meteorit. Planet. Sci.* **41**, 95–106 (2006).
- Haack, H., Rasmussen, K. L. & Warren, P. H. Effects of regolith/megaregolith insulation on the cooling histories of differentiated asteroids. *J. Geophys. Res.* **95**, 5111–5124 (1990).
- Warren, P. H. Ejecta-megaregolith accumulation on planetesimals and large asteroids. *Meteorit. Planet. Sci.* **46**, 53–78 (2011).
- Sahijpal, S., Soni, P. & Gupta, G. Numerical simulations of the differentiation of accreting planetesimals with  $^{26}\text{Al}$  and  $^{60}\text{Fe}$  as the heat sources. *Meteorit. Planet. Sci.* **42**, 1529–1548 (2007).



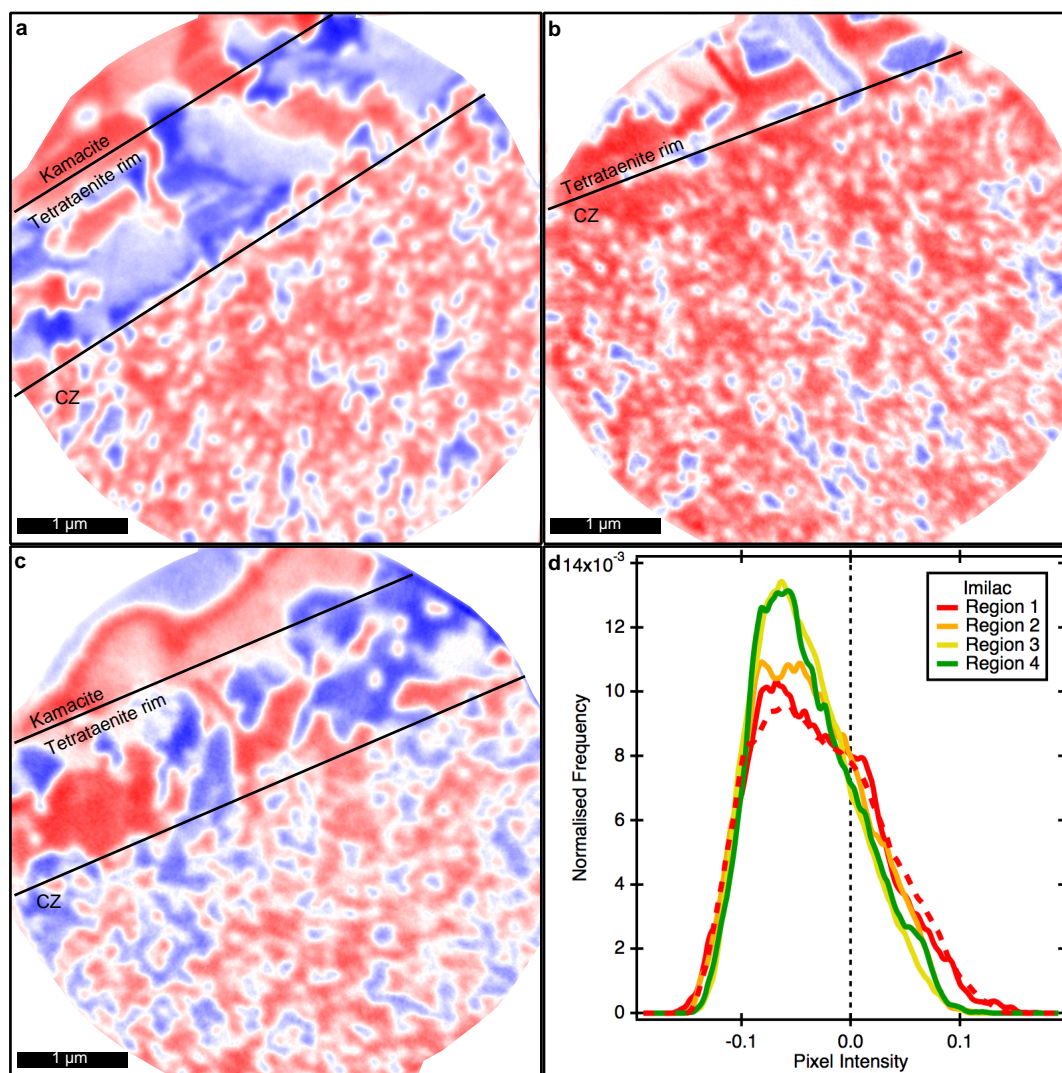
**Extended Data Figure 1 | A 15- $\mu\text{m}$  field-of-view XPEEM image of the kamacite, tetrataenite rim, CZ and aligned CZ in the Imilac meteorite.** The boundaries between the kamacite and tetrataenite rim and the tetrataenite rim and CZ are marked as solid black lines, and the boundary between the CZ and aligned CZ is marked as a dashed line. Note that the interface between the CZ and fine aligned CZ is abrupt and parallel to the tetrataenite rim. The in-plane orientation of the X-ray beam is included as an arrow; the beam was  $16^\circ$  out of plane. This orientation applies to all images.



**Extended Data Figure 2 | Representative SEM images of the CZ in the pallasites.** Coarser tetraenaite islands in the Imilac (a) and Esquel (b) meteorites. The regions corresponding to Fig. 1 are included (red boxes).

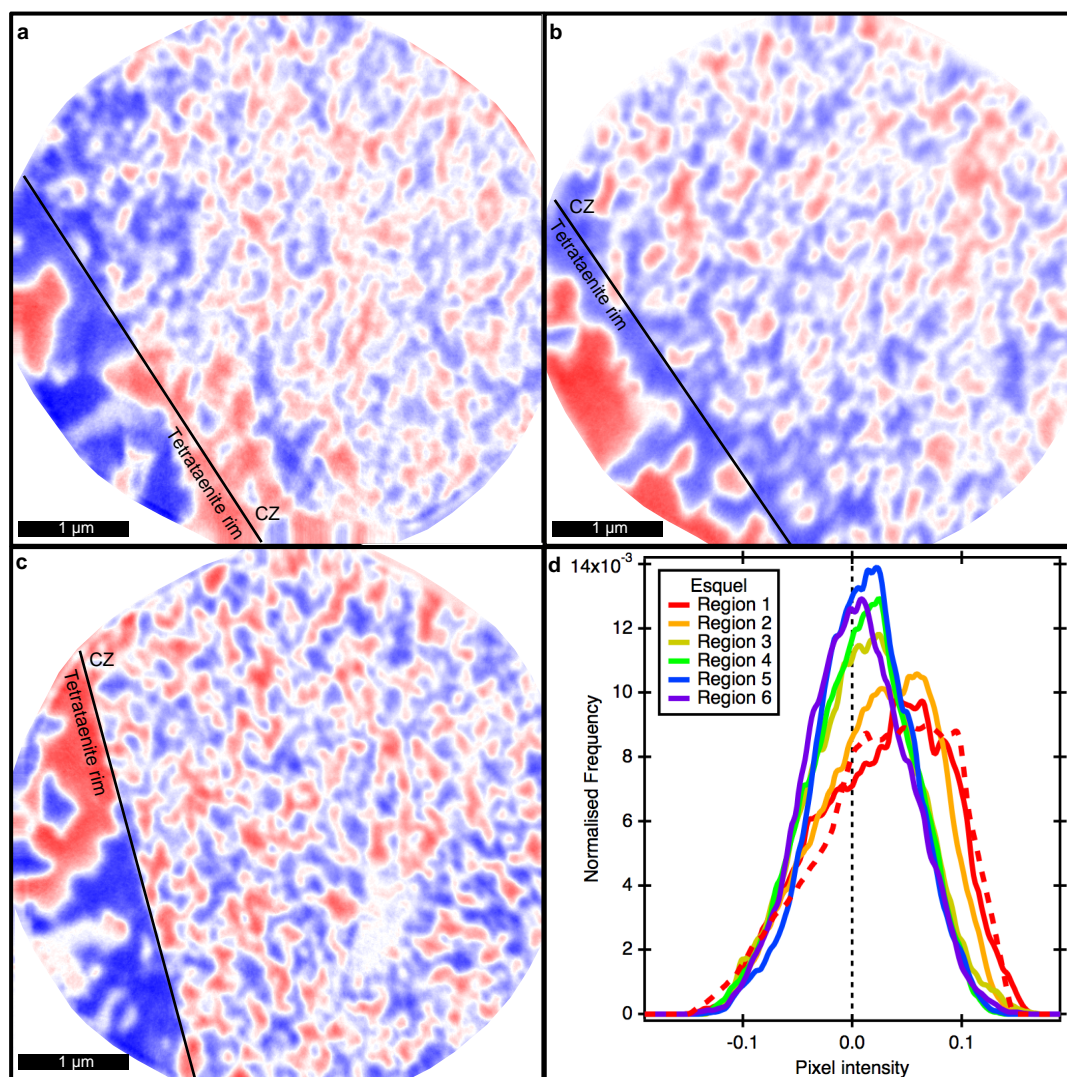
The CZ is unclear and island sizes could not be extracted beyond four and six regions of the Imilac and Esquel meteorites, respectively. The average island size of each region is included in Extended Data Table 1.



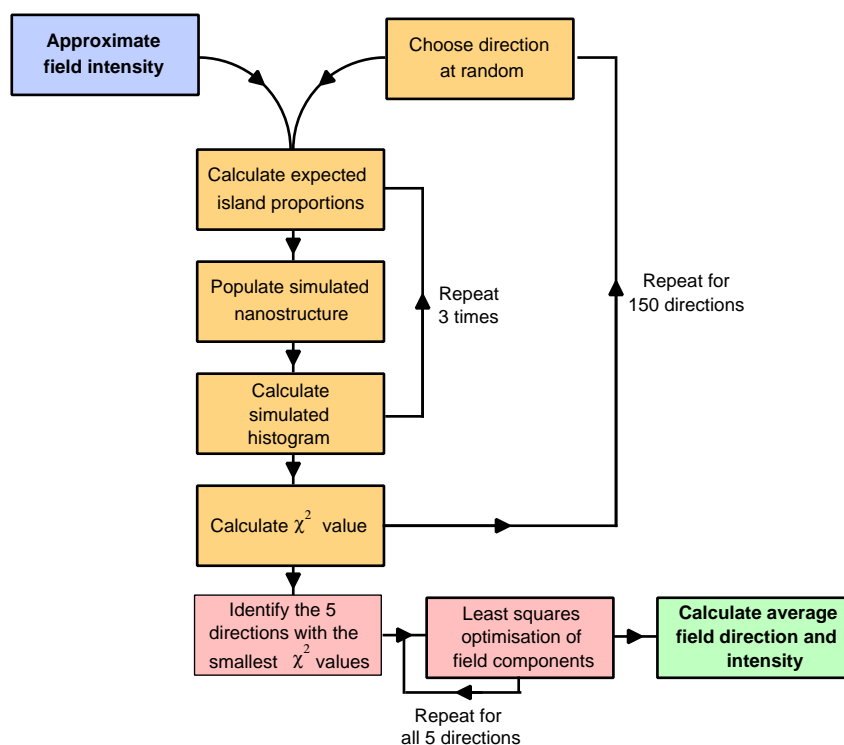


**Extended Data Figure 3 | Additional XPEEM images and pixel intensity histograms for the Imilac meteorite.** a–c, 5-μm field-of-view XPEEM images from three areas of the Imilac meteorite. These areas, along with Fig. 1a, passed the selection criteria outlined in the ‘Experimental analysis’ section of the Methods. Note that the magnetic patterns vary between images. d, Average

pixel intensity histograms calculated from the four images of the Imilac meteorite. The dashed line is the optimized fit curve for the best fit for region 1; the fits for other regions were not included (to reduce clutter), but all displayed similar agreement to their experimental counterpart.



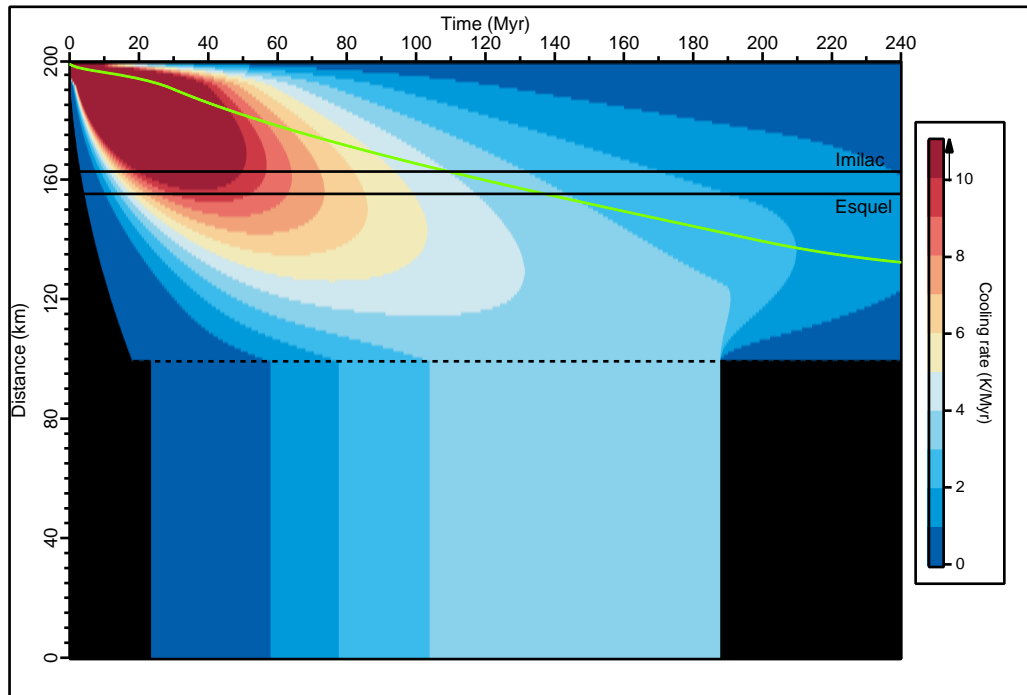
Extended Data Figure 4 | As for Extended Data Fig. 3 but for the Esquel meteorite.



**Extended Data Figure 5 | Flow diagram of the procedure used to identify the intensity and orientation of the field recorded by each region, as described in the Methods.** The blue box represents the input to the procedure (approximate field intensity). The orange boxes represent the parts of the

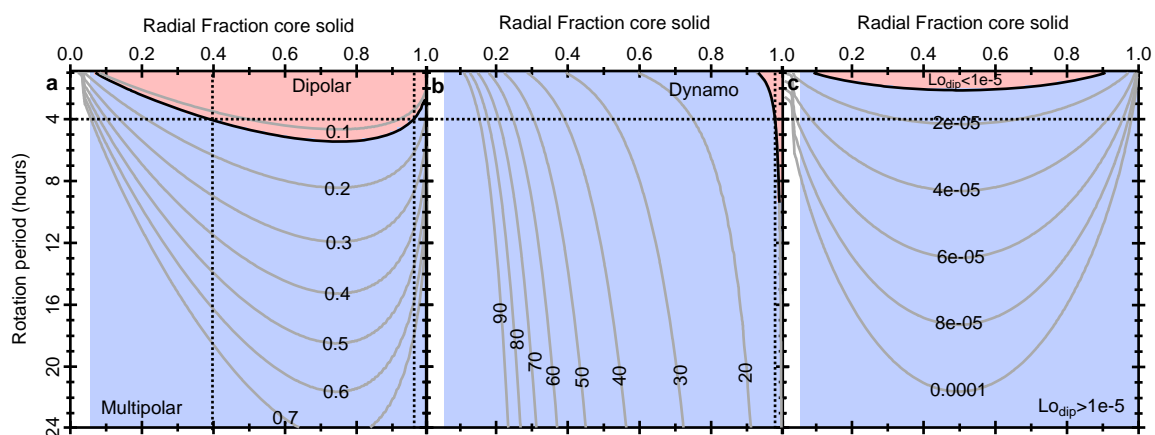
procedure used to deduce the approximate direction of the recorded field. The red boxes represent the part of the procedure used to optimize the direction and intensity estimates. The green box represents the final outputs.





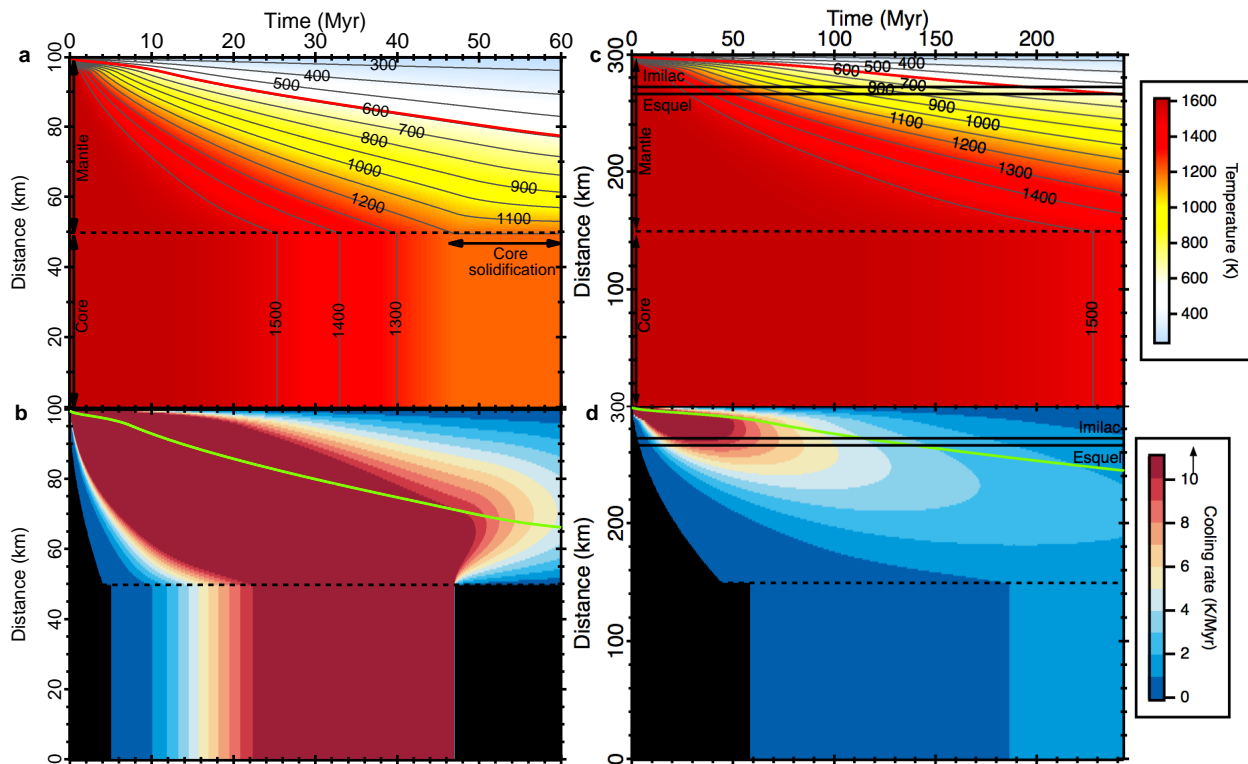
**Extended Data Figure 6 | 200-km-radius body cooling rate.** Values derived from Fig. 4. Cooling rates between  $0 \text{ K Myr}^{-1}$  and  $10 \text{ K Myr}^{-1}$  are depicted by the colour scale; any value greater than  $10 \text{ K Myr}^{-1}$  is depicted by the dark

red colour. The green line is the 800-K contour from Fig. 4, which corresponds to the temperature of the cooling rates inferred from the relationship to island size<sup>14</sup>. The core cools at a rate of between  $0 \text{ K Myr}^{-1}$  and  $3.6 \text{ K Myr}^{-1}$ .



**Extended Data Figure 7 | Planetary magnetic parameters calculated from the dynamo generation model.** **a**, Local Rossby number,  $Ro_l$ , showing the combination of parameters resulting in predominantly dipolar (red region) to multipolar (blue region) fields. **b**, Magnetic Reynolds number,  $Re_m$ , showing

the combination of parameters that do (blue region) and do not (red region) result in dynamo activity. **c**, Dipolar Lorentz number,  $Lo_{dip}$ , showing the combination of parameters that do (blue region) and do not (red region) produce  $Lo_{dip} > 10^{-5}$ . A four-hour rotation period is highlighted in each figure.



#### Extended Data Figure 8 | Additional planetary cooling models.

**a, b,** Temperatures (**a**) and cooling rates (**b**) from the conductive cooling model with a 100-km radius. The cooling rate at 800 K (green line) reaches the inferred values of the Imilac or the Esquel meteorites after complete core solidification.

**c, d,** Temperatures (**c**) and cooling rate (**d**) from the conductive cooling model with a 300-km radius. The depths (horizontal black lines) corresponding to the inferred cooling rates at 800 K (green line) for the Imilac and Esquel meteorites reach the tetrataenite formation temperature (593 K, red line) before the core solidification starts.



**Extended Data Table 1 | Average island sizes for each region of the Imilac and Esquel meteorite studied**

Region number	Imilac meteorite (nm)	Esquel meteorite (nm)
Region 1	130	154
Region 2	111	125
Region 3	95	113
Region 4	87	97
Region 5	-	85
Region 6	-	75

Values extracted from Extended Data Fig. 2. There are only four regions for the Imilac meteorite.

Extended Data Table 2 | Values of the parameters used in the planetary cooling and dynamo generation models

Parameter	Value	Unit
$\kappa$	$5 \times 10^{-7}$	$\text{m}^2 \text{s}^{-1}$
$\delta r$	1000	m
$\delta t$	$2 \times 10^{11}$	s
$r_c$	100000	m
$k_m$	3	$\text{W m}^{-1} \text{K}^{-1}$
$\rho_c$	7800	$\text{kg}^3 \text{m}^{-3}$
$C_p$	850	$\text{J kg}^{-1} \text{K}^{-1}$
$L_c$	270000	$\text{J kg}^{-1}$

All values are supported by experimental and/or theoretical work<sup>1,7,35</sup>. From the top row to the bottom, the parameters are the mantle thermal diffusivity, the incremental distance, the incremental time, the core radius, the mantle thermal conductivity, the density of the core material, the heat capacity of the core material and the latent heat of core solidification.

# Compositional engineering of perovskite materials for high-performance solar cells

Nam Joong Jeon<sup>1\*</sup>, Jun Hong Noh<sup>1\*</sup>, Woon Seok Yang<sup>1</sup>, Young Chan Kim<sup>1</sup>, Seungchan Ryu<sup>1</sup>, Jangwon Seo<sup>1</sup> & Sang Il Seok<sup>1,2</sup>

**Of the many materials and methodologies aimed at producing low-cost, efficient photovoltaic cells, inorganic–organic lead halide perovskite materials<sup>1–17</sup> appear particularly promising for next-generation solar devices owing to their high power conversion efficiency. The highest efficiencies reported for perovskite solar cells so far have been obtained mainly with methylammonium lead halide materials<sup>1–10</sup>. Here we combine the promising—owing to its comparatively narrow bandgap—but relatively unstable formamidinium lead iodide (FAPbI<sub>3</sub>) with methylammonium lead bromide (MAPbBr<sub>3</sub>) as the light-harvesting unit in a bilayer solar-cell architecture<sup>13</sup>. We investigated phase stability, morphology of the perovskite layer, hysteresis in current–voltage characteristics, and overall performance as a function of chemical composition. Our results show that incorporation of MAPbBr<sub>3</sub> into FAPbI<sub>3</sub> stabilizes the perovskite phase of FAPbI<sub>3</sub> and improves the power conversion efficiency of the solar cell to more than 18 per cent under a standard illumination of 100 milliwatts per square centimetre. These findings further emphasize the versatility and performance potential of inorganic–organic lead halide perovskite materials for photovoltaic applications.**

An inorganic–organic lead halide perovskite is any material that crystallizes into an AMX<sub>3</sub> (where A is an organic ammonium cation, M is Pb or Sn, and X is a halide anion) structure. The size of cation A is critical for the formation of a close-packed perovskite structure; in particular, cation A must fit into the space composed of four adjacent corner-sharing MX<sub>6</sub> octahedra. Of the various inorganic–organic lead halide perovskite materials, methylammonium lead iodide (MAPbI<sub>3</sub>), with a bandgap of about 1.5–1.6 eV and a light absorption spectrum up to a wavelength of 800 nm, has been extensively used as a light harvester in solar cells. Several methods, including a one-step spin-coating method<sup>1–3,6</sup>, a two-step sequential method<sup>4</sup>, and vapour deposition in a high-vacuum chamber<sup>5</sup>, have been used to prepare MAPbI<sub>3</sub> materials. Additionally, different device architectures such as planar<sup>5,10</sup> and mesostructured cells<sup>1,7</sup> have also been proposed. The highest power conversion efficiency<sup>8,9,11</sup> (PCE) so far, of 16% to 17%, was achieved for solution-processed MAPbI<sub>3</sub>, although 19.3% in a planar device architecture was recently obtained from a reverse-bias current–voltage (*I*–*V*) curve<sup>12</sup>. We note that a PCE measured<sup>13–15</sup> via a reverse-bias scan with the solar cells exhibiting a large-hysteresis *I*–*V* curve can be highly overestimated.

We have already showed that chemical modification of the X site anions (for example, substitution of I with Br) of MAPbI<sub>3</sub> can tune the bandgaps to range between 1.5 eV and 2.3 eV by incorporating MAPbBr<sub>3</sub> (2.3 eV bandgap); this resulted in colour variation and PCE modulation<sup>6</sup>. In contrast, HC(NH<sub>2</sub>)<sub>2</sub>PbI<sub>3</sub>, which contains formamidinium (FA) cations instead of methylammonium (MA) cations at the ‘A’ site of the AMX<sub>3</sub> perovskite structure, has a bandgap of 1.48 eV, with an absorption edge of 840 nm (refs 16–18). We expect that this reduced bandgap may allow absorption of photons over a broader solar spectrum. The structural and opto-electrical differences of MAPbI<sub>3</sub> and FAPbI<sub>3</sub> are likely to originate from the difference in ionic radius of the MA (1.8 Å) and FA ions (1.9–2.2 Å). In fact, the relative ionic radii of A, M and X in the

AMX<sub>3</sub> perovskite structure have been widely used as a method of establishing the distortion of the MX<sub>6</sub> octahedron; in particular, a relatively smaller ion radius for X favours the formation of cubic structures<sup>18</sup>. However, the photovoltaic performance of FAPbI<sub>3</sub> has been reported to be lower than that of MAPbI<sub>3</sub> (refs 16, 17). In addition, the black perovskite-type polymorph ( $\alpha$ -phase), which is stable at relatively high temperatures (above 160 °C), was observed to turn into the yellow FAPbI<sub>3</sub> polymorph ( $\delta$ -phase) in an ambient humid atmosphere<sup>17</sup>. However, considering their suitable bandgap (which is lower than that in MAPbI<sub>3</sub>), the performance of FAPbI<sub>3</sub> solar cells can be considerably improved by stabilizing the FAPbI<sub>3</sub> phase, improving the crystallinity, and optimizing the cell architecture. In this regard, Pellet *et al.*<sup>19</sup> demonstrated an improved PCE using mixed cation lead iodide perovskites by gradually substituting MA with FA cations—which increases the absorption range by shifting it redwards, allowing for a higher current density—but the performance was still dominated by MAPbI<sub>3</sub> rather than FAPbI<sub>3</sub>. We have recently reported a 16.2% certified PCE obtained from the combination of MAPbI<sub>3</sub> and MAPbBr<sub>3</sub> with a bilayer architecture consisting of perovskite-infiltrated mesoporous-TiO<sub>2</sub> electrodes, and an extremely uniform and dense upper perovskite layer obtained by solvent engineering techniques<sup>13</sup>; the absorption edge was below 770 nm. A strategy for extending the absorption range of solar light is thus to replace MAPbI<sub>3</sub> with FAPbI<sub>3</sub> in the combined composition of MAPbI<sub>3</sub> and MAPbBr<sub>3</sub>.

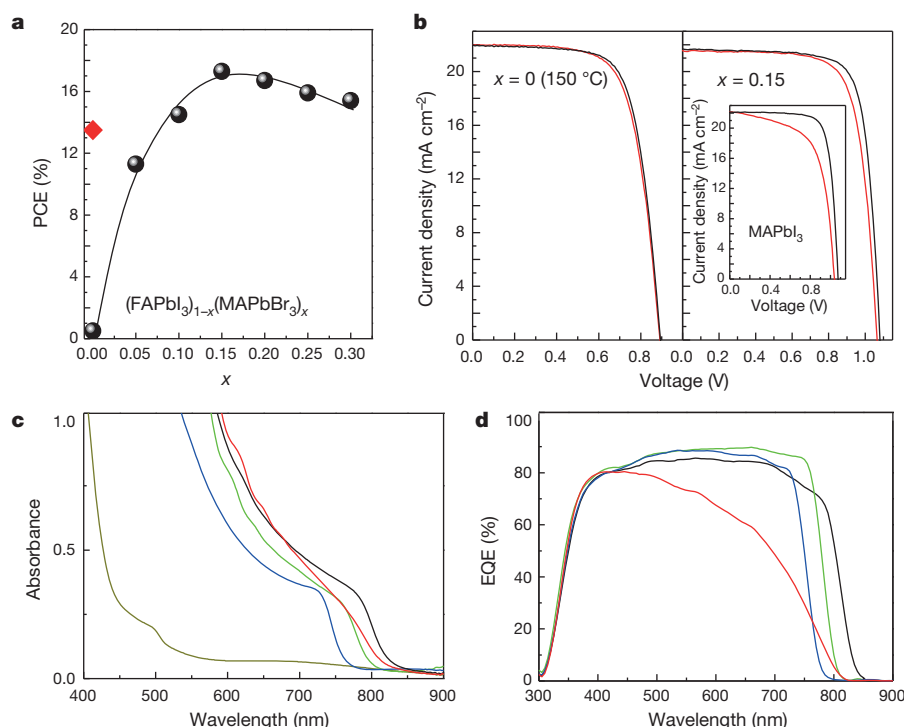
Here we report the overall PCE and structural stability of (FAPbI<sub>3</sub>)<sub>1–x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub>, with the mole ratio *x* ranging between 0 and 0.3. We investigated the performance of systems containing different amounts of MAPbBr<sub>3</sub> incorporated into FAPbI<sub>3</sub>—that is, (FAPbI<sub>3</sub>)<sub>1–x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub>—by measuring their current density versus voltage (*J*–*V*). Figure 1a shows the PCEs obtained from the *J*–*V* curves for different (FAPbI<sub>3</sub>)<sub>1–x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub> systems under standard air-mass 1.5 global (AM1.5G) illumination. The PCEs were averaged from the reverse (from the open-circuit voltage *V*<sub>oc</sub> to the short-circuit current *I*<sub>sc</sub>) and forward (from *I*<sub>sc</sub> to *V*<sub>oc</sub>) sweeps, because perovskite solar cells may experience a dependence of the measured efficiency on the scan direction. However, the hysteresis of the cells fabricated from FAPbI<sub>3</sub> and FAPbI<sub>3</sub> incorporating 15 mol% MAPbBr<sub>3</sub> was not relevant, even after collecting *J*–*V* curves with the relatively short scanning delay time of 40 ms (Fig. 1b). In addition, the small discrepancies related to the scan direction for the FAPbI<sub>3</sub>/MAPbBr<sub>3</sub> systems disappeared at longer delay times (above 100 ms). That the hysteresis observed in FAPbI<sub>3</sub>-based systems is negligible compared to the hysteresis in MAPbI<sub>3</sub> (Fig. 1b inset) may be related to the balance between electron and hole transport within the perovskite layer.

As reported, the electron-diffusion length in MAPbI<sub>3</sub> (~130 nm) is 1.4 times larger than the hole-diffusion length (~90 nm)<sup>20</sup>, while FAPbI<sub>3</sub> has a hole-diffusion length (~813 nm) that is 4.6 times longer than the electron-diffusion length (~177 nm)<sup>16</sup>. In addition, Kanatzidis *et al.*<sup>18</sup> showed, by measuring the Seebeck coefficient, that MAPbI<sub>3</sub> and FAPbI<sub>3</sub> display an n-type and a p-type character, respectively. Since the incident light of the FTO/TiO<sub>2</sub>/perovskite/PTAA/Au (where FTO is

<sup>1</sup>Division of Advanced Materials, Korea Research Institute of Chemical Technology, 141 Gajeong-Ro, Yuseong-Gu, Daejeon 305-600, South Korea. <sup>2</sup>Department of Energy Science, 2066 Seoburo, Jangang-gu, Sungkyunkwan University, Suwon 440-746, South Korea.

\*These authors contributed equally to this work.





**Figure 1 | Characterization of materials.** **a**, PCE values for cells using  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  materials, which were annealed at  $100^\circ\text{C}$  for 10 min (black line); the red diamond indicates the PCE value for the cell fabricated using pure  $\text{FAPbI}_3$  ( $x = 0$ ); annealing was performed at  $150^\circ\text{C}$  for 10 min to form the black perovskite phase. **b**,  $J-V$  curves for cells constructed from  $\text{FAPbI}_3$ ,  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$  and  $\text{MAPbI}_3$  (inset) films measured with a 40 ms scanning delay in reverse (from  $V_{\text{oc}}$  to  $I_{\text{sc}}$ ) and forward (from  $I_{\text{sc}}$  to  $V_{\text{oc}}$ ) modes under standard AM 1.5 G illumination. Pure  $\text{FAPbI}_3$  ( $x = 0$ ) was then annealed at  $150^\circ\text{C}$ . **c**, Ultraviolet–visible absorption spectra of  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films (with  $x = 0$ , dark yellow;  $x = 0.05$ , red;  $x = 0.15$ , green;  $x = 0.25$ , blue) annealed at  $100^\circ\text{C}$  on fused silica substrates; the black line indicates absorption spectra for the pure  $\text{FAPbI}_3$  film (perovskite phase) annealed at  $150^\circ\text{C}$ ; the dark yellow line indicates a non-perovskite  $\text{FAPbI}_3$  form. **d**, EQE spectra for cells using  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  materials (with  $x = 0$ , black;  $x = 0.05$ , red;  $x = 0.15$ , green;  $x = 0.25$ , blue). Pure  $\text{FAPbI}_3$  ( $x = 0$ ) was annealed at  $150^\circ\text{C}$ .

fluorine-doped tin oxide and PTAA is poly(triarylamine)) cell configuration reaches the perovskite through the FTO side, more carriers are generated on the perovskite near  $\text{TiO}_2$  than on the perovskite near PTAA. Therefore, many more holes than electrons should travel a long distance through the perovskite layer to the PTAA layer. This indicates that the balance between electron and hole transport in cells should be considered in compositional engineering for highly efficient perovskite solar cells. Data in Table 1 show that the solar cell annealed at  $100^\circ\text{C}$  with  $x = 0$  exhibits considerably low  $J_{\text{sc}}$ ,  $V_{\text{oc}}$  and fill factor (FF) values. In contrast, the overall PCE, including  $J_{\text{sc}}$ ,  $V_{\text{oc}}$  and FF, noticeably increased when annealing was performed at  $150^\circ\text{C}$ . The trend of PCE as a function of  $x$  in the cells fabricated with  $\text{FAPbI}_3$ – $\text{MAPbBr}_3$  showed that  $J_{\text{sc}}$  increases from  $19.0 \text{ mA cm}^{-2}$  (at  $x = 0.05$ ) to a maximum value of  $22 \text{ mA cm}^{-2}$  (at  $x = 0.15$ ); after this point,  $J_{\text{sc}}$  decreases to  $20 \text{ mA cm}^{-2}$  (at  $x = 0.30$ ). In this range,  $V_{\text{oc}}$  increases from  $1.0 \text{ V}$  to  $1.12 \text{ V}$ . Interestingly, the FF showed exactly the same  $J_{\text{sc}}$  behaviour with a maximum value of 73% at  $x = 0.15$ .

The ultraviolet–visible absorption spectra displayed in Fig. 1c show a systematic shift of the absorption band edge to shorter wavelengths when the  $\text{MAPbBr}_3$  content is increased. The reduction of  $J_{\text{sc}}$  observed with  $x$  larger than 0.15 is directly related to the blue-shift of absorption onset, which in turn is responsible for the reduced light-harvesting efficiency. As expected, the external quantum efficiency (EQE) spectrum shown in Fig. 1d is blue-shifted when  $x$  is increased, resulting in the reduction of  $J_{\text{sc}}$ . However, a relatively lower  $J_{\text{sc}}$ , below  $x = 0.15$ , indicates that the charge-collection efficiency is also low, because  $J_{\text{sc}}$  is

proportional to the product of the charge-collection efficiency and the light-harvesting efficiency. The fact that the FF has a trend similar to that of  $J_{\text{sc}}$  supports this, as confirmed by the considerably high series resistance shown below  $x = 0.15$  (with annealing at  $100^\circ\text{C}$ ). The increase of  $V_{\text{oc}}$  over the entire range studied in this work may be attributed to the widening of the bandgap, which occurs when  $x$  increases. Most importantly, the PCE of the solar cells fabricated in this study exhibits a maximum value of 17.3% at  $x = 0.15$ , owing to the simultaneous enhancement of  $J_{\text{sc}}$  and FF, while  $V_{\text{oc}}$  continues to increase (because the PCE is determined from the products of  $J_{\text{sc}}$ ,  $V_{\text{oc}}$  and FF).

To further elucidate possible reasons for such low values of  $J_{\text{sc}}$  and FF below  $x = 0.15$ , we examined the phase stability of  $\text{FAPbI}_3$ . It is known that  $\text{FAPbI}_3$  is characterized by a reversible phase transition between two polymorphs, that is, a black perovskite with trigonal symmetry (space group  $P3m1$ ) and a yellow non-perovskite with hexagonal symmetry ( $P6_3mc$ )<sup>17,18</sup>. The latter contains linear chains of  $[\text{PbI}_6]$  octahedrons with face-sharing, while the former consists of a three-dimensional network of corner-sharing octahedrons.

Figure 2a shows the results of the differential scanning calorimetry (DSC) and thermogravimetric analysis of a yellow  $\text{FAPbI}_3$  powder heated at  $300^\circ\text{C}$ . An endothermic peak around  $160^\circ\text{C}$  was found with the DSC analysis, which appears without any weight loss up to  $250^\circ\text{C}$  in the thermogravimetric analysis. The X-ray diffraction (XRD) spectra of the yellow powder measured *in situ* (Extended Data Fig. 1) suggest that the endothermic peak can be assigned to the phase transition from yellow non-perovskite to black perovskite, in agreement with previous studies<sup>16,21</sup>. This indicates that the yellow non-perovskite phase for the pure  $\text{FAPbI}_3$  material is likely to be thermodynamically stable. This phase transition was also found to be reversible in air, that is, the prepared yellow non-perovskite phase changed to the black perovskite phase when annealing was performed at  $170^\circ\text{C}$ ; the black powder turned yellow again after being stored in air for 10 days (Extended Data Fig. 2).

This unintentional phase transition in  $\text{FAPbI}_3$  solar cells may reduce the photovoltaic performance, because the yellow phase is characterized by a large optical bandgap (Fig. 1c) and an inferior charge-transporting ability due to the linear chain-like  $[\text{PbI}_6]$  octahedron structure. Notably,  $\text{AMX}_3$  (where A is Rb, Cs, MA or FA, M is Pb or Sn, and X is Cl, Br or I) metal trihalide materials exist as either two polymorphs (perovskite and non-perovskite) or only one of the two, depending on the atomic size of

**Table 1 | Photovoltaic parameters from  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$**

$x$	$J_{\text{sc}}$ ( $\text{mA cm}^{-2}$ )	$V_{\text{oc}}$ (V)	FF	PCE (%)	Series resistance ( $\Omega \text{ cm}^2$ )
0 ( $150^\circ\text{C}$ )	22.0	0.88	0.70	13.5	5.7
0	1.10	0.88	0.51	0.5	345
0.05	17.1	1.02	0.65	11.3	6.0
0.10	21.0	1.04	0.66	14.5	4.8
0.15	22.0	1.08	0.73	17.3	3.9
0.20	21.5	1.09	0.71	16.7	4.3
0.25	21.0	1.10	0.69	15.9	4.9
0.30	20.0	1.12	0.69	15.4	5.7

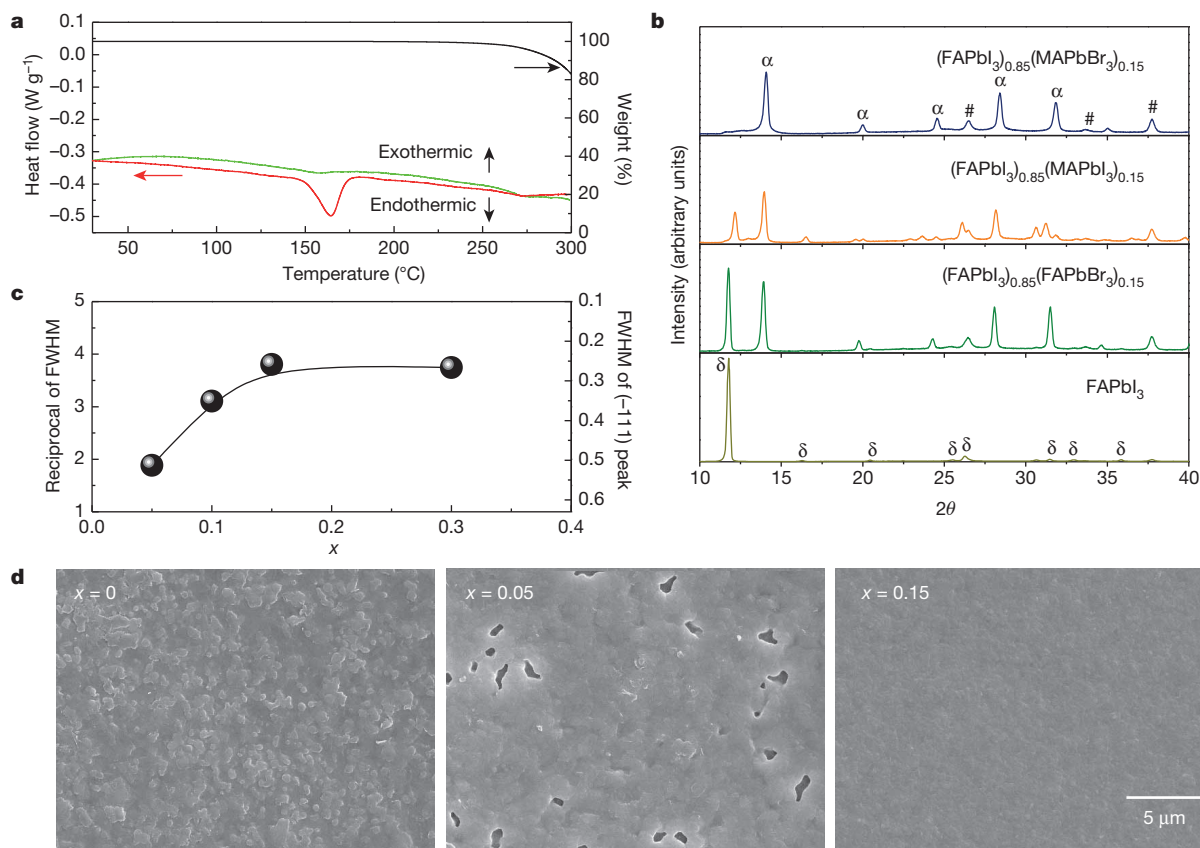
the components<sup>22</sup>. For example, in contrast to  $\text{CsSnI}_3$ , which has two polymorphs,  $\text{MASnI}_3$ , with larger  $\text{MA}^+$  cations, shows only a perovskite phase and no phase transition near room temperature ( $\sim 25^\circ\text{C}$ ).  $\text{RbSnI}_3$ , with smaller  $\text{Rb}^+$  cations, shows only a non-perovskite phase<sup>22–24</sup>. Therefore, the atomic combination of the A and X sites may lead to the stabilization of the  $\text{FAPbI}_3$  perovskite structure.

Figure 2b shows the XRD spectra of the prepared  $\text{FAPbI}_3$ ,  $(\text{FAPbI}_3)_{1-x}(\text{MAPbI}_3)_x$ ,  $(\text{FAPbI}_3)_{1-x}(\text{FAPbBr}_3)_x$  and  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films (with  $x = 0.15$ ) on mesoporous- $\text{TiO}_2$ /blocking- $\text{TiO}_2$ /FTO glass substrates, after annealing at  $100^\circ\text{C}$  for 10 min; these were prepared using the solvent-engineering process with precursor solutions of the desired compositions as previously reported<sup>13</sup>. The XRD spectrum of the pure  $\text{FAPbI}_3$  thin film shows the typical diffraction pattern of a hexagonal non-perovskite polymorph of  $\text{FAPbI}_3$  ( $P6_3mc$ )<sup>18</sup>; this can be detected from DSC (Fig. 2a), because the temperature of the annealing ( $100^\circ\text{C}$ ) is much lower than the temperature ( $160^\circ\text{C}$ ) at which the phase transition occurs. However, when  $\text{FA}^+$  cations in  $\text{FAPbI}_3$  are substituted by 15 mol% of  $\text{MA}^+$  cations, a strong (111) diffraction peak at  $13.9^\circ$  for the trigonal perovskite phase ( $P3m1$ ) appears in spite of the annealing at  $100^\circ\text{C}$ . The same diffraction peaks are also observed in systems containing  $\text{Br}^-$  ions (15 mol%), although the secondary phase coexists in the film.

Surprisingly, a simultaneous introduction of 15 mol% of both  $\text{MA}^+$  cations and  $\text{Br}^-$  anions in  $\text{FAPbI}_3$  to obtain  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$  leads to a synergetic effect that stabilizes the perovskite phase. Interestingly, this is sufficient to form a  $\text{FAPbI}_3$  perovskite phase even after incorporating 5 mol% of  $\text{MAPbBr}_3$  (Extended Data Fig. 3), although

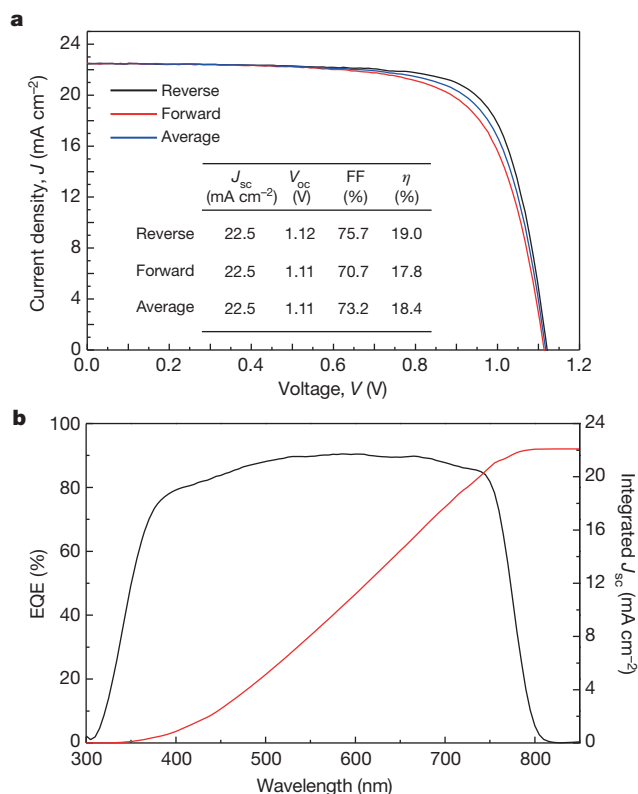
the single  $\text{MA}^+$  or  $\text{Br}^-$  can only partially form a perovskite phase. We also found that a highly crystalline perovskite layer is formed with values of  $x$  larger than 0.15, according to the full width of half maximum (FWHM) of the (–111) diffraction peak (Fig. 2c). Therefore, we surmise that the enhancement of the phase stability and crystallinity results in an improvement of PCE in the  $x$  range of 0 to 0.15. The perovskite phase stabilization caused by the introduction of  $\text{MAPbBr}_3$  to  $\text{FAPbI}_3$  was also confirmed by an analysis carried out on the synthesized powders, which were prepared at room temperature by precipitation from solutions of  $\text{FAPbI}_3$ ,  $(\text{FAPbI}_3)_{1-x}(\text{MAPbI}_3)_x$ ,  $(\text{FAPbI}_3)_{1-x}(\text{FAPbBr}_3)_x$  and  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  with  $x = 0.15$ . Photographs of the as-prepared powders shown in Extended Data Fig. 4 indicate that a black powder like as-prepared  $\text{MAPbI}_3$  is obtained only for  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$ . In addition, the XRD spectra of the powders revealed that  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$ , in contrast to  $\text{FAPbI}_3$ , shows a pure perovskite phase (Extended Data Fig. 5), with no endothermic DSC peaks (Fig. 2a). This finding confirms that the co-substitution of MA to FA and Br to I can efficiently stabilize the perovskite phase. However, further investigation is required to determine the energetics of the perovskite and non-perovskite formation and to establish the composition of the stable form in perovskite halide materials.

We recently successfully fabricated a complete bilayer on a mesoporous- $\text{TiO}_2$  using the solvent-engineering technology; our findings confirmed that surface coverage and morphology of perovskite materials is critical<sup>13</sup>. For this reason, here we analysed the surfaces deposited with different  $x$  values for  $\text{MAPbBr}_3/\text{FAPbI}_3$  using the solvent-engineering process. Figure 2d shows the surface scanning electron microscope (SEM) images



**Figure 2 | Characterization of materials.** **a**, DSC and thermogravimetric curves of the as-prepared yellow  $\text{FAPbI}_3$  powder under an Ar atmosphere with a heating rate of  $2^\circ\text{C min}^{-1}$  from room temperature to  $300^\circ\text{C}$ . The green line indicates the DSC results for  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$  powder, which was prepared at room temperature by precipitation from the same compositional solution. **b**, XRD spectra of solvent-engineering-processed  $\text{FAPbI}_3$ ,  $(\text{FAPbI}_3)_{1-x}(\text{MAPbI}_3)_x$ ,  $(\text{FAPbI}_3)_{1-x}(\text{FAPbBr}_3)_x$  and  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films with  $x = 0.15$ , on mesoporous- $\text{TiO}_2$ /

blocking- $\text{TiO}_2$ /FTO glass substrates after annealing at  $100^\circ\text{C}$  for 10 min.  $\alpha$ ,  $\delta$  and  $\#$  denote the identified diffraction peaks corresponding to the perovskite and non-perovskite polymorphs of  $\text{FAPbI}_3$  and FTO, respectively. **c**, FWHM of the (–111) peak for  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films as a function of  $x$  from XRD spectra of solvent-engineering-processed  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films (see Extended Data Fig. 4 for further details). **d**, SEM plane view images of  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  films with  $x = 0, 0.05$  and  $0.15$ .



**Figure 3** | *J*-*V* and IPCE characteristics for the best cell obtained in this study. **a**, *J*-*V* curves of forward and reverse bias sweep and their averaged curve for the solar cell using the (FAPbI<sub>3</sub>)<sub>0.85</sub>(MAPbBr<sub>3</sub>)<sub>0.15</sub> perovskite active layer. **b**, EQE spectrum and integrated  $J_{sc}$ .

of (FAPbI<sub>3</sub>)<sub>1-x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub>, with  $x = 0, 0.05$  and  $0.15$ , on a 200-nm-thick mesoporous-TiO<sub>2</sub>/blocking-TiO<sub>2</sub>/FTO annealed for 10 min at 150 °C ( $x = 0$ ) and 100 °C, respectively. Our results showed that the surface of FAPbI<sub>3</sub> exhibits an irregular morphology with bumpy roughness. Incorporating MAPbBr<sub>3</sub> into FAPbI<sub>3</sub> (with  $x = 0.15$ ) considerably smoothed the surface; however, systems characterized by  $x = 0.05$  still presented large voids between crystal boundaries. The rough surface of FAPbI<sub>3</sub> may be due to the transition from non-perovskite to perovskite phases and to the high temperature required for the formation of the perovskite phase. Indeed, manipulating the composition of FAPbI<sub>3</sub> by adding MAPbBr<sub>3</sub> led to the stabilization of the perovskite phase with a uniform and dense morphology as well as well-developed crystallites, which are responsible for the highly improved cell performance.

On the basis of these results, we repeated the fabrication procedure, fixing  $x$  to a value of 0.15 to enhance further the performance of the FAPbI<sub>3</sub>-MAPbBr<sub>3</sub> system with an architecture as follows: FTO/blocking-TiO<sub>2</sub> (70 nm)/mesoporous-TiO<sub>2</sub>/perovskite composite layer (200 nm)/perovskite upper layer (300 nm)/PTAA (50 nm)/Au (100 nm). Of the devices obtained, Fig. 3a shows the *J*-*V* curves measured via reverse and forward bias sweep for one of the best-performing solar cells. The  $J_{sc}$ ,  $V_{oc}$  and FF values averaged from the *J*-*V* curves of this device are 22.5 mA cm<sup>-2</sup>, 1,105 mV and 73.2%, respectively; these correspond to a PCE of 18.4% under standard AM1.5G conditions. The PCE value is in agreement with that obtained from the stabilized power output near the maximum power point, reflecting the device performance in the working condition more closely (Extended Data Fig. 6).

In the case of the cell using thinner mesoporous-TiO<sub>2</sub> layer (80 nm), as shown in Extended Data Fig. 7, although an unprecedented PCE of 20.3% was measured via reverse bias scan, the PCE of around 17.3% obtained from an average *J*-*V* curve and steady-state current measurement is far lower than that from the best cell using a 200-nm-thick mesoporous-TiO<sub>2</sub> layer, owing to the low PCE of 15.5% with forward

bias scan. This result is similar to those of ref. 12 and implies that substantial PCE values should be obtained from *J*-*V* curves averaged with reverse and forward bias sweep. Figure 3b shows the EQE spectrum for one of the best-performing solar cells. A very broad EQE plateau of over 80% between 400 nm and 750 nm was observed. The  $J_{sc}$  value integrated from EQE was found to be in good agreement with that measured by *J*-*V*. The highly performing devices exhibiting PCEs of 18.0% with very small hysteresis was certified by the standardized method in the photovoltaic calibration laboratory, confirming a PCE of 17.9% under AM1.5G full sun (Extended Data Fig. 8). To the best of our knowledge, such a PCE value is the highest ever reported for a perovskite-based solar cell, excluding the value overestimated by reverse bias scan.

We have described the compositional engineering of (FAPbI<sub>3</sub>)<sub>1-x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub> for efficient perovskite solar cells. (FAPbI<sub>3</sub>)<sub>0.85</sub>(MAPbBr<sub>3</sub>)<sub>0.15</sub> has many advantages over other systems such as MAPbI<sub>3</sub>, FAPbI<sub>3</sub>, and MAPb(I<sub>0.85</sub>Br<sub>0.15</sub>)<sub>3</sub>. This strategy may lead to more efficient and cost-effective inorganic-organic hybrid perovskite solar cells.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 October; accepted 2 December 2014.

Published online 7 January 2015.

- Lee, M. M. *et al.* Efficient hybrid solar cells based on meso-structured organometal halide perovskites. *Science* **338**, 643–647 (2012).
- Kim, H.-S. *et al.* Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%. *Sci. Rep.* **2**, 1–7 (2012).
- Heo, J. H. *et al.* Efficient inorganic-organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors. *Nature Photon.* **7**, 486–491 (2013).
- Burschka, J. *et al.* Sequential deposition as a route to high-performance perovskite-sensitized solar cells. *Nature* **499**, 316–319 (2013).
- Liu, M., Johnston, M. B. & Snaith, H. J. Efficient planar heterojunction perovskite solar cells by vapor deposition. *Nature* **501**, 395–398 (2013).
- Noh, J. H., Im, S. H., Heo, J. H., Mandal, T. N. & Seok, S. I. Chemical management for colorful, efficient, and stable inorganic-organic hybrid nanostructured solar cells. *Nano Lett.* **13**, 1764–1769 (2013).
- Ball, J. M., Lee, M. M., Hey, A. & Snaith, H. J. Low-temperature processed meso-structured to thin-film solar cells. *Energy Environ. Sci.* **6**, 1739–1743 (2013).
- Jeon, N. J. *et al.* *o*-Methoxy substituents in Spiro-OMeTAD for efficient inorganic-organic hybrid perovskite solar cells. *J. Am. Chem. Soc.* **136**, 7837–7840 (2014).
- Ryu, S. *et al.* Voltage output of efficient perovskite solar cells with high open-circuit voltage and fill factor. *Energy Environ. Sci.* **7**, 2614–2618 (2014).
- Malinkiewicz, O. *et al.* Perovskite solar cells employing organic charge-transport layers. *Nature Photon.* **8**, 128–132 (2014).
- Lee, J.-W. *et al.* High-efficiency perovskite solar cells based on the black polymorph of HC(NH<sub>2</sub>)<sub>2</sub>PbI<sub>3</sub>. *Adv. Mater.* **26**, 4991–4998 (2014).
- Zhou, H. *et al.* Interface engineering of highly efficient perovskite solar cells. *Science* **345**, 542–546 (2014).
- Jeon, N. J. *et al.* Solvent engineering for high-performance inorganic-organic hybrid perovskite solar cells. *Nature Mater.* **13**, 897–903 (2014).
- Snaith, H. J. *et al.* Anomalous hysteresis in perovskite solar cells. *J. Phys. Chem. Lett.* **5**, 1511–1515 (2014).
- Kim, H. S. & Park, N.-G. Parameters affecting I-V hysteresis of CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub> perovskite solar cells: effect of perovskite crystal size and mesoporous TiO<sub>2</sub> layer. *J. Phys. Chem. Lett.* **5**, 2927–2934 (2014).
- Eperon, G. E. *et al.* Formamidinium lead halide: a broad tunable perovskite for efficient planar heterojunction solar cells. *Energy Environ. Sci.* **7**, 982–988 (2014).
- Koh, T. M. *et al.* Formamidinium-containing metal-halide: an alternative material for near-IR absorption perovskite solar cells. *J. Phys. Chem. C* **118**, 16458–16462 (2014).
- Stoumpos, C. C. *et al.* Semiconducting tin and lead iodide perovskites with organic cations: phase transition, high mobilities, and near-infrared photoluminescent properties. *Inorg. Chem.* **52**, 9019–9038 (2013).
- Pellet, N. *et al.* Mixed-organic-cation perovskite photovoltaics for enhanced solar-light harvesting. *Angew. Chem. Int. Ed.* **53**, 3151–3157 (2014).
- Xing, G. *et al.* Long-range balanced electron and hole-transport lengths in organic-inorganic CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub>. *Science* **342**, 344–347 (2013).
- Scaife, D. E., Weller, P. F. & Fisher, W. G. Crystal preparation and properties of cesium tin(II) trihalides. *J. Solid State Chem.* **9**, 308–314 (1974).
- Chung, I. *et al.* CsSnI<sub>3</sub>: semiconductor or metal? High electrical conductivity and strong near-infrared photoluminescence from a single material. High hole mobility and phase-transitions. *J. Am. Chem. Soc.* **134**, 8579–8587 (2012).
- Takahashi, Y. *et al.* Charge-transport in tin-iodide perovskite CH<sub>3</sub>NH<sub>3</sub>SnI<sub>3</sub>: origin of high conductivity. *Dalton Trans.* **50**, 5563–5568 (2011).
- Amat, A. *et al.* Cation-induced band-gap tuning in organohalide perovskites: interplay of spin-orbit coupling and octahedra tilting. *Nano Lett.* **14**, 3608–3616 (2014).



**Acknowledgements** This work was supported by the Global Research Laboratory (GRL) Program, the Global Frontier R&D Program of the Center for Multiscale Energy System, funded by the National Research Foundation in Korea, and by a grant from the Korea Research Institute of Chemical Technology (KRICT) 2020 Program for Future Technology in South Korea.

**Author Contributions** N.J.J., J.H.N. and S.I.S. conceived the experiments and analysed and interpreted the data. N.J.J., Y.C.K., J.H.N. and J.S. performed the fabrication of devices, device performance measurements and characterization. N.J.J., W.S.Y. and S.R. carried out the synthesis of materials for perovskites, and

S.I.S. prepared  $\text{TiO}_2$  particles and pastes. The manuscript was mainly written and revised by S.I.S. and J.H.N. The project was planned, directed and supervised by S.I.S. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.I.S. ([seoksi@kRICT.re.kr](mailto:seoksi@kRICT.re.kr) or [seoksi@skku.edu](mailto:seoksi@skku.edu)).

## METHODS

**Materials.** Unless stated otherwise, all materials were purchased from Sigma-Aldrich or Junsei Organics and used as received. PTAA was purchased from EM Index.

**Synthesis of the TiO<sub>2</sub> paste.** The TiO<sub>2</sub> nanoparticles used in the paste were prepared by hydrothermal treatment at 250 °C for 12 h from aqueous solutions of the peroxotitanium complex, as described elsewhere<sup>25,26</sup>. The peroxotitanium complex solutions were synthesized via a reaction between hydrogen peroxide and the TiO(OH)<sub>2</sub> wet cake. The wet cake was obtained by the dropwise addition of a NH<sub>4</sub>OH solution into 0.33M TiOCl<sub>2</sub> aqueous solution obtained by the hydrolysis of TiCl<sub>4</sub>. The resulting cake was purified by repeated washing with deionized water until no Cl<sup>−</sup> ions were detected. Then, 200 ml of 35 wt% hydrogen peroxide was added into 300 ml of a 2 wt% TiO(OH)<sub>2</sub>-dispersed aqueous solution such that the H<sub>2</sub>O<sub>2</sub>/Ti<sup>4+</sup> ratio was ~40; the solution was constantly stirred in an ice bath. Once the addition of H<sub>2</sub>O<sub>2</sub> was completed, the colour of the precipitate changed from white to turbid yellowish, and finally light orange- or reddish-coloured transparent solutions were formed within 1 h.

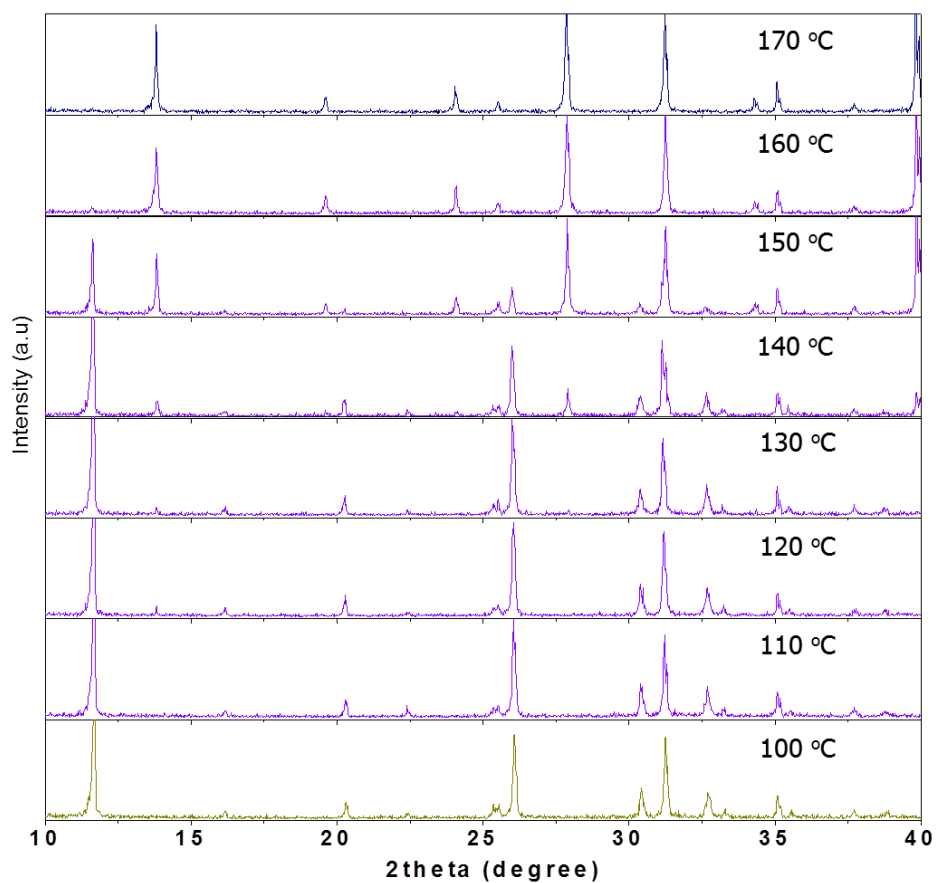
The TiO<sub>2</sub>-nanoparticle (average diameter 50 nm, anatase) slurry obtained by the hydrothermal treatment was stirred for 1 h, while several drops of concentrated nitric acid were added to obtain a colloidal dispersion solution. Then the TiO<sub>2</sub> nanoparticles were collected by the centrifugal method. The TiO<sub>2</sub> wet cake was re-dispersed and centrifuged in absolute ethanol by ultrasonic irradiation. The process with absolute ethanol was repeated for three cycles. The collected TiO<sub>2</sub> nanoparticles were dispersed in 100 ml of absolute ethanol, 4.5 g (per 1 g of TiO<sub>2</sub>) of a 10 wt% ethanolic solution of ethyl cellulose (Fluka, 46070), and 4.4 g (per 1 g of TiO<sub>2</sub>) of terpineol. After the addition of each component, the mix was stirred for 10 min and homogenized by ultrasonic irradiation. The paste was finally produced by a three-roller-mill grinder (EXAKT) after concentrating the mixture solution in a rotary evaporator. **Synthesis of the inorganic–organic mixed perovskite.** CH<sub>3</sub>NH<sub>3</sub>I and NH<sub>2</sub>CH = NH<sub>2</sub>I were first synthesized by reacting 30 ml hydroiodic acid (57% in water, Aldrich), 27.86 ml CH<sub>3</sub>NH<sub>2</sub> (40% in methanol, Junsei Chemical), and 15 g formamidine acetate (Aldrich) in a 250 ml round-bottomed flask at 0 °C for 2 h with stirring. The precipitates were recovered by evaporating the solutions at 50 °C for 1 h. The products were dissolved in ethanol, recrystallized using diethyl ether, and finally dried at 60 °C in a vacuum oven for 24 h. Similarly, CH<sub>3</sub>NH<sub>3</sub>Br and NH<sub>2</sub>CH = NH<sub>2</sub>Br were prepared using hydrobromic acid (48 wt% in water, Aldrich) according to a reported procedure<sup>27</sup>. The desired solutions of FAPbI<sub>3</sub>, (FAPbI<sub>3</sub>)<sub>1−x</sub>(MAPbI<sub>3</sub>)<sub>x</sub>, (FAPbI<sub>3</sub>)<sub>1−x</sub>(FAPbBr<sub>3</sub>)<sub>x</sub>, and (FAPbI<sub>3</sub>)<sub>1−x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub> (with  $x = 0–0.30$ ) were prepared by dissolution of the CH<sub>3</sub>NH<sub>3</sub>I, CH<sub>3</sub>NH<sub>3</sub>Br, NH<sub>2</sub>CH = NH<sub>2</sub>I, and NH<sub>2</sub>CH = NH<sub>2</sub>Br powders with PbI<sub>2</sub> (Aldrich) and PbBr<sub>2</sub> (Aldrich) in the  $\gamma$ -butyrolactone:DMSO mixed solvent (7:3, volume ratio) at 60 °C for 10 min.

**Solar cell fabrication.** A dense blocking layer of TiO<sub>2</sub> (60 nm, blocking-TiO<sub>2</sub>) was deposited onto an F-doped SnO<sub>2</sub> (FTO, Pilkington, TEC8) substrate by spray pyrolysis using a 20 mM titanium diisopropoxide bis(acetylacetonate) solution (Aldrich) at 450 °C; this was done to prevent a direct contact between FTO and the hole-conducting layer. A 200-nm-thick mesoporous-TiO<sub>2</sub> was spin-coated onto the blocking-TiO<sub>2</sub>/FTO substrate using TiO<sub>2</sub> pastes diluted in 2-methoxyethanol (1 g in 5 ml) and calcinated at 500 °C for 1 h in air to remove the organic components.

The inorganic–organic lead halide perovskite solutions were then coated onto the mesoporous-TiO<sub>2</sub>/blocking-TiO<sub>2</sub>/FTO substrate by two consecutive spin-coating steps, at 1,000 rpm and 5,000 rpm for 40 s and 20 s, respectively. During the second spin-coating step, 1 ml toluene was poured onto the substrate, according to the procedure reported in ref. 13. To obtain a uniform and flat intermediate-phase film, toluene should be quickly cast in one-shot mode on a rapidly rotating (at 5,000 rpm) substrate to wash out the surplus of DMSO molecules that did not participate in the formation of the PbI<sub>2</sub>–NH<sub>2</sub>CH = NH<sub>2</sub>I–DMSO complex. The substrate was then dried on a hot plate at 100 °C or 150 °C for 10 min. A solution of PTAA (EM Index, [Mn] = 17,500 g mol<sup>−1</sup>)/toluene (10 mg ml<sup>−1</sup>) with an additive of 7.5  $\mu$ l Li-bis(trifluoromethanesulphonyl) imide/acetonitrile (170 mg ml<sup>−1</sup>) and 4  $\mu$ l 4-*tert*-butylpyridine was spin-coated on the perovskite layer/mesoporous-TiO<sub>2</sub>/blocking-TiO<sub>2</sub>/FTO substrate at 3,000 rpm for 30 s. Finally, an Au counter electrode was deposited by thermal evaporation; the active area of this electrode was fixed at 0.16 cm<sup>2</sup>. The different inorganic–organic lead halide triiodide powders were prepared by precipitation using toluene at room temperature from the desired compositional solutions obtained by dissolving CH<sub>3</sub>NH<sub>3</sub>I, CH<sub>3</sub>NH<sub>3</sub>Br, NH<sub>2</sub>CH = NH<sub>2</sub>I and NH<sub>2</sub>CH = NH<sub>2</sub>Br powders with PbI<sub>2</sub> and PbBr<sub>2</sub> in  $\gamma$ -butyrolactone.

**Characterization.** The XRD spectra of the prepared films were measured using a Rigaku SmartLab X-ray diffractometer; the *in situ* XRD experiment of the as-prepared FAPbI<sub>3</sub> yellow powder was performed using a Rigaku Ultima IV with an X-ray tube (Cu K $\alpha$ , wavelength  $\lambda = 1.5406$  Å). Ultraviolet–visible absorption spectra were recorded on a Shimadzu UV 2550 spectrophotometer in the 300–800 nm wavelength range at room temperature. The morphology of the films was observed using a field-emission SEM (MIRA3 LMU, Tescan). Thermogravimetric and DSC analyses of the as-prepared powders were performed with a heating rate of 2 °C min<sup>−1</sup> from room temperature up to 300 °C under a nitrogen atmosphere using TA Instruments SDT 2960 and DSC 2910, respectively. EQE was measured by a power source (Newport 300W Xenon lamp, 66920) with a monochromator (Newport Cornerstone 260) and a multimeter (Keithley2001). The  $J$ – $V$  curves were measured using a solar simulator (Newport, Oriel Class A, 91195A) with a source meter (Keithley 2420) at 100 mA cm<sup>−2</sup> AM1.5G illumination and a calibrated Si-reference cell certified by the National Renewable Energy Laboratory, USA. The  $J$ – $V$  curves were measured by reverse scan (forward bias (1.2 V)  $\rightarrow$  short circuit (0 V)) or forward scan (short circuit (0 V)  $\rightarrow$  forward bias (1.2 V)). The step voltage was fixed at 10 mV and the delay time, which is a delay set at each voltage step before measuring each current, was modulated. The  $J$ – $V$  curves for all devices were measured by masking the active area with a metal mask (area of 0.096 cm<sup>2</sup>).

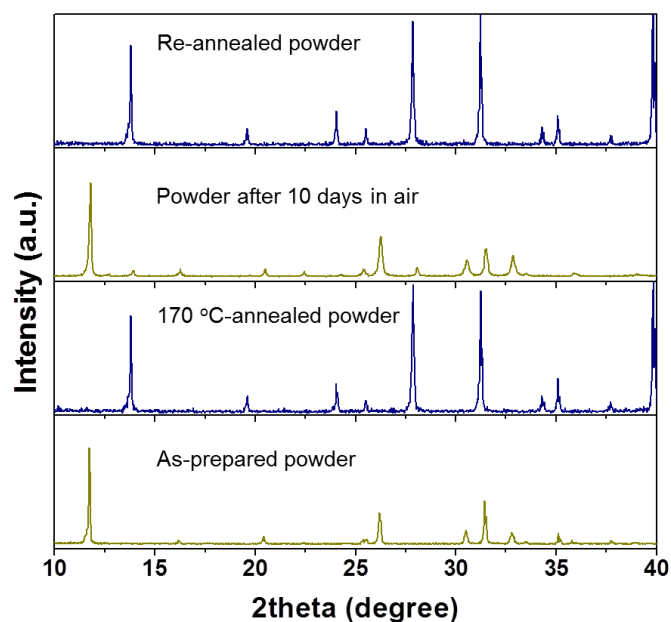
25. Baek, I. C. *et al.* Facile preparation of large aspect ratio ellipsoidal anatase TiO<sub>2</sub> nanoparticles and their application to dye-sensitized solar cell. *Electrochem. Commun.* **11**, 909–912 (2009).
26. Seok, S. I. *et al.* Colloidal TiO<sub>2</sub> nanocrystals prepared from peroxotitanium complex solutions: phase evolution from different precursors. *J. Colloid Interf. Sci.* **346**, 66–71 (2010).
27. Pang, S. *et al.* NH<sub>2</sub>CH=NH<sub>2</sub>PbI<sub>3</sub>: an alternative organolead iodide perovskite sensitizer for mesoscopic solar cells. *Chem. Mater.* **26**, 1485–1491 (2014).



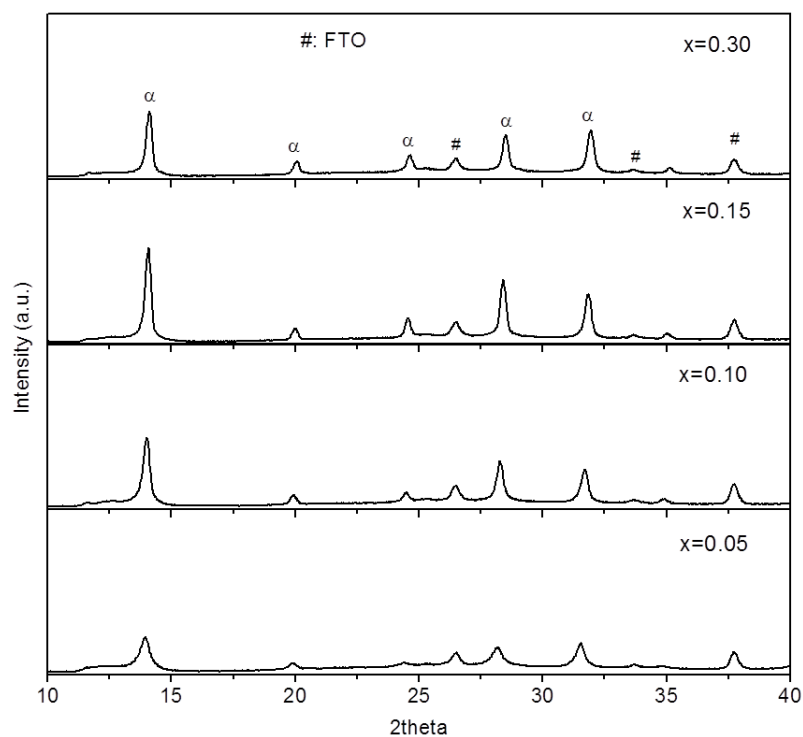
**Extended Data Figure 1 | *In situ* XRD spectra (heating from 100 °C to 170 °C) for FAPbI<sub>3</sub> yellow powders prepared at room temperature. Hexagonal non-perovskite FAPbI<sub>3</sub> (*P*<sub>6</sub><sub>3</sub>*mc*) converted into a trigonal**

**perovskite phase (*P*<sub>3</sub>*m*<sub>1</sub>) near 150 °C. The (−111) diffraction peak for perovskite FAPbI<sub>3</sub> at  $2\theta = 14.3^\circ$  appeared at a temperature of 150 °C; simultaneously the main peak of non-perovskite FAPbI<sub>3</sub> at  $11.6^\circ$  disappeared.**



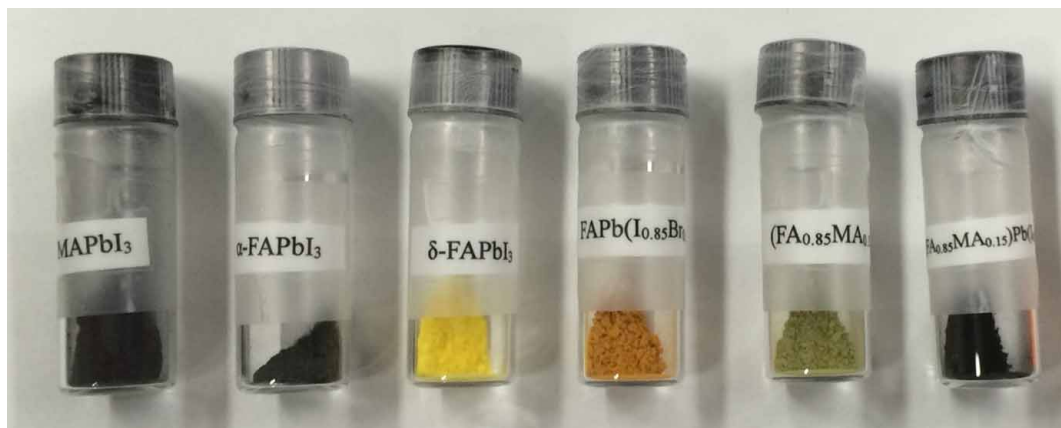


**Extended Data Figure 2 | XRD spectra of FAPbI<sub>3</sub> powders.** The as-prepared yellow FAPbI<sub>3</sub> powder shows a non-perovskite phase and is converted to perovskite phase by annealing at 170 °C. The perovskite FAPbI<sub>3</sub> black powder returned to the yellow non-perovskite powder after being stored in air for 10 h; the yellow powder reversibly changed to black perovskite phase by re-annealing at 170 °C.



**Extended Data Figure 3 | XRD spectra of  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  cells as a function of  $x$ .** XRD spectra of solvent-engineering processed  $(\text{FA}_{1-x}\text{MA}_x)\text{Pb}(\text{I}_{1-x}\text{Br}_x)_3$  films on the mesoporous- $\text{TiO}_2$ /blocking- $\text{TiO}_2$ /

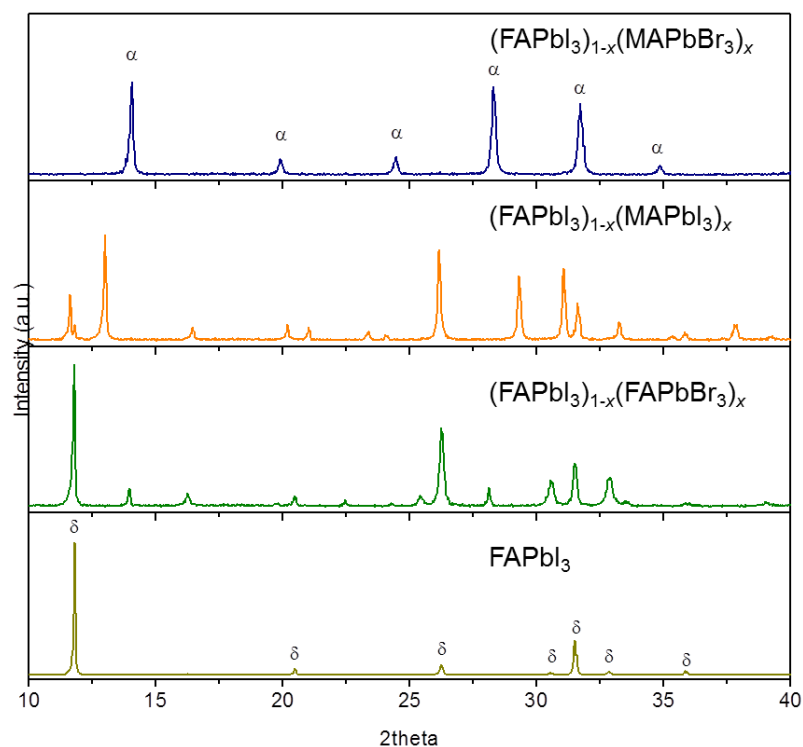
FTO glass substrate after annealing at  $100^\circ\text{C}$  for 10 min.  $\alpha$ ,  $\alpha$ -phase of FAPbI<sub>3</sub>; #, peaks diffracted from FTO.



**Extended Data Figure 4 | Photographs of inorganic-organic hybrid halide powders.** Photographs show the colour of the as-prepared MAPbI<sub>3</sub>, annealed FAPbI<sub>3</sub> at 170 °C, FAPbI<sub>3</sub>, (FAPbI<sub>3</sub>)<sub>1-x</sub>(MAPbI<sub>3</sub>)<sub>x</sub>,

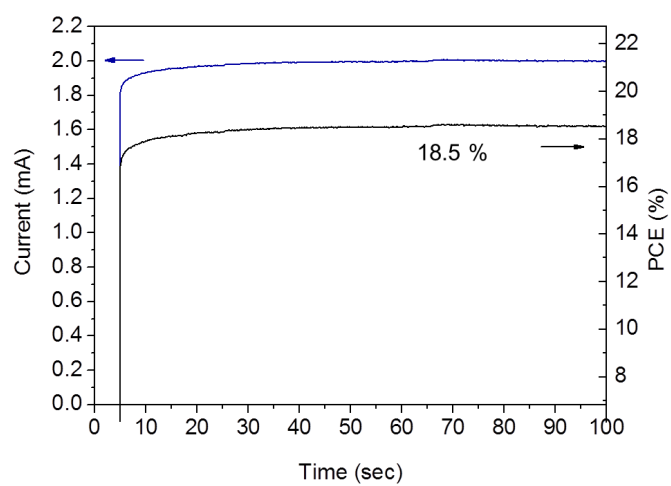
(FAPbI<sub>3</sub>)<sub>1-x</sub>(FAPbBr<sub>3</sub>)<sub>x</sub>, and (FAPbI<sub>3</sub>)<sub>1-x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub> powders with  $x = 0.15$  (from left to right). The (FAPbI<sub>3</sub>)<sub>1-x</sub>(MAPbBr<sub>3</sub>)<sub>x</sub> powder is the only black powder among the as-prepared FAPbI<sub>3</sub>-based materials.



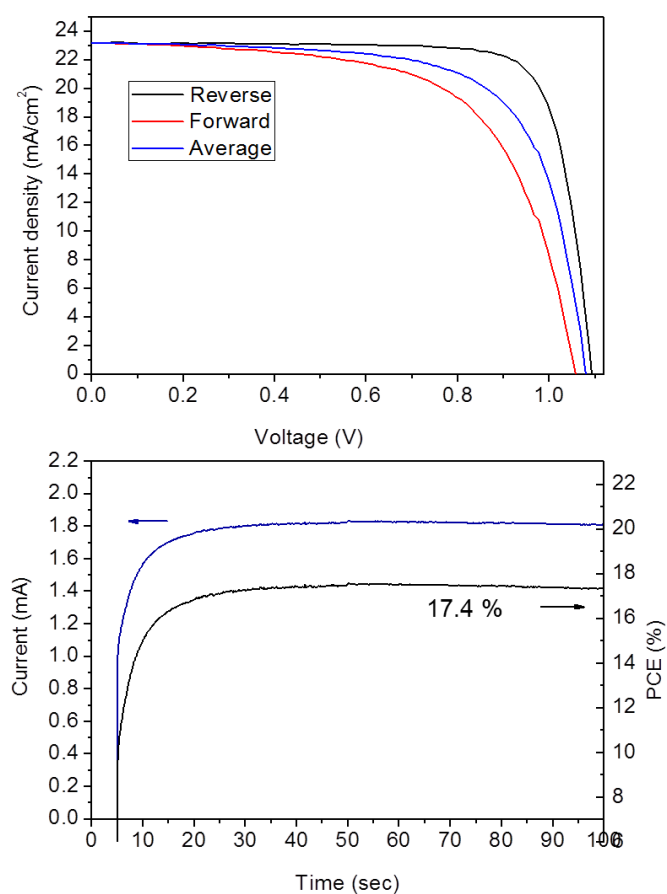


**Extended Data Figure 5 | XRD spectra of the as-prepared powders at room temperature.** XRD spectra of the as-prepared  $\text{FAPbI}_3$ ,  $(\text{FAPbI}_3)_{1-x}(\text{MAPbI}_3)_x$ ,  $(\text{FAPbI}_3)_{1-x}(\text{FAPbBr}_3)_x$ , and

$(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  powders with  $x = 0.15$  (from left to right). Only the  $(\text{FAPbI}_3)_{1-x}(\text{MAPbBr}_3)_x$  powder shows a pure perovskite phase.  $\alpha$ , black perovskite-type polymorph;  $\delta$ , yellow non-perovskite polymorph.

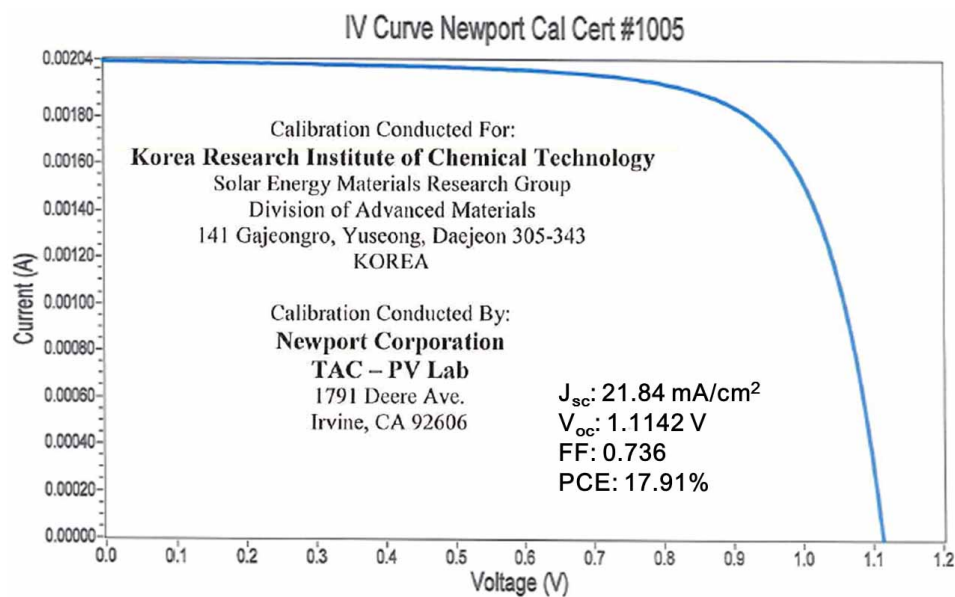


**Extended Data Figure 6 | Steady-state current measurement.** Steady-state current measured at a maximum power point (0.89 V) and stabilized power output.



**Extended Data Figure 7 | Photovoltaic performance.** **a**,  $J$ - $V$  curves measured by forward and reverse bias sweep and their averaged curve for cell using the  $(\text{FAPbI}_3)_{0.85}(\text{MAPbBr}_3)_{0.15}$  perovskite active layer and 80-nm-thick mesoporous- $\text{TiO}_2$  layer.  $\eta$ , PCE. **b**, Steady-state current measured at a maximum power point (0.92 V) and stabilized power output.





Extended Data Figure 8 | Independent certification from Newport Corporation, confirming a PCE of 17.9%.

# Probabilistic reanalysis of twentieth-century sea-level rise

Carling C. Hay<sup>1,2</sup>, Eric Morrow<sup>1,2</sup>, Robert E. Kopp<sup>2,3</sup> & Jerry X. Mitrovica<sup>1</sup>

Estimating and accounting for twentieth-century global mean sea-level (GMSL) rise is critical to characterizing current and future human-induced sea-level change. Several previous analyses of tide gauge records<sup>1–6</sup>—employing different methods to accommodate the spatial sparsity and temporal incompleteness of the data and to constrain the geometry of long-term sea-level change—have concluded that GMSL rose over the twentieth century at a mean rate of 1.6 to 1.9 millimetres per year. Efforts to account for this rate by summing estimates of individual contributions from glacier and ice-sheet mass loss, ocean thermal expansion, and changes in land water storage fall significantly short in the period before 1990<sup>7</sup>. The failure to close the budget of GMSL during this period has led to suggestions that several contributions may have been systematically underestimated<sup>8</sup>. However, the extent to which the limitations of tide gauge analyses have affected estimates of the GMSL rate of change is unclear. Here we revisit estimates of twentieth-century GMSL rise using probabilistic techniques<sup>9,10</sup> and find a rate of GMSL rise from 1901 to 1990 of  $1.2 \pm 0.2$  millimetres per year (90% confidence interval). Based on individual contributions tabulated in the Fifth Assessment Report<sup>7</sup> of the Intergovernmental Panel on Climate Change, this estimate closes the twentieth-century sea-level budget. Our analysis, which combines tide gauge records with physics-based and model-derived geometries of the various contributing signals, also indicates that GMSL rose at a rate of  $3.0 \pm 0.7$  millimetres per year between 1993 and 2010, consistent with prior estimates from tide gauge records<sup>4</sup>. The increase in rate relative to the 1901–90 trend is accordingly larger than previously thought; this revision may affect some projections<sup>11</sup> of future sea-level rise.

Tide gauges provide records of local sea-level changes that, in the case of some sites, extend back to the eighteenth century<sup>12–14</sup>. However, using the database of tide gauge records<sup>15</sup> to estimate historical GMSL rise (defined as the increase in ocean volume normalized by ocean area) is challenging. Tide gauges sample the ocean sparsely and non-uniformly, with a bias towards coastal sites and the Northern Hemisphere, and with few sites at latitudes greater than  $60^\circ$  (see, for example, refs 4, 9). In addition, tide gauge time series show significant inter-annual to decadal variability, and they are characterized by missing data (that is, intervals without observations at the start, middle or end of a time series). From the perspective of estimating GMSL changes, the data are contaminated by local and regional signals due to ongoing glacial isostatic adjustment (GIA) associated with past ice ages<sup>16,17</sup>, the spatially non-uniform pattern of sea-level rise associated with changes in contemporary land ice sources<sup>18–21</sup>, ocean/atmosphere dynamics<sup>22</sup>, and other local factors including tectonics, sediment compaction, groundwater pumping and harbour development.

Different approaches have been used to address these complexities in efforts to estimate twentieth-century GMSL rise<sup>23</sup>. These include averaging rates at sites with the longest records<sup>1,2</sup>, averaging rates determined from regional binning of records<sup>3</sup>, incorporating shorter records into the analysis to distinguish between secular trends and decadal-scale variability<sup>3</sup>, and using altimetry records to determine dominant

sea-level geometries and then using tide gauge records to estimate the time-varying amplitudes of these geometries<sup>4,5</sup>. In most cases, other criteria were applied to cull the tide gauge sites adopted in the analysis (for example, excluding sites near tectonic activity or major urban centres).

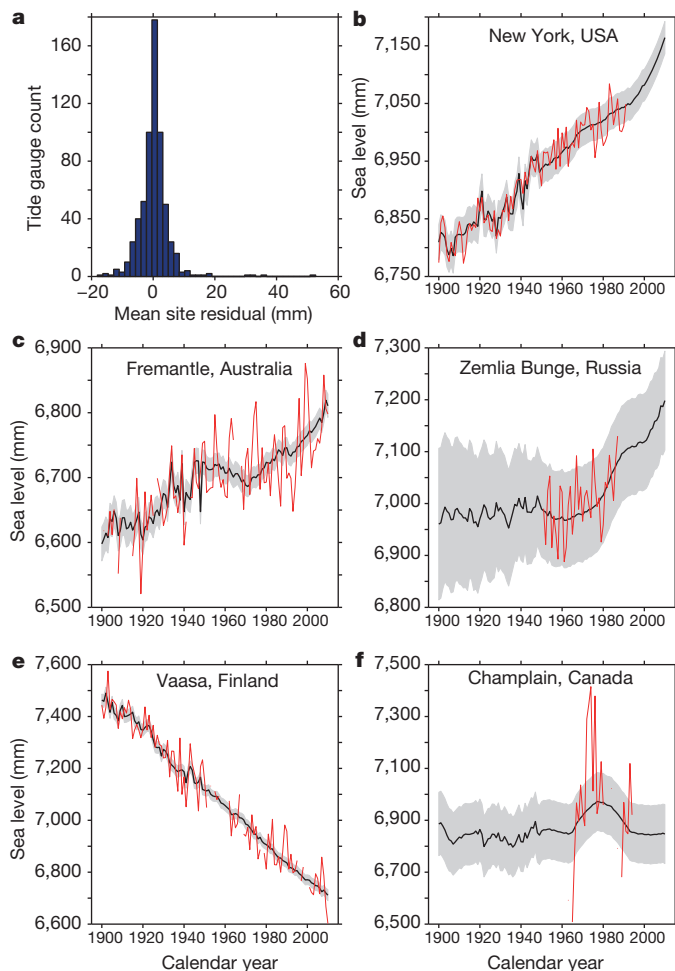
Estimates of twentieth-century GMSL rise from these previous analyses range from 1.6 to 1.9  $\text{mm yr}^{-1}$  (refs 1–6) and define an important enigma. Independent model- and data-based estimates of the individual sources of GMSL, including mass flux from glaciers and ice sheets, thermal expansion of oceans, and changes in land water storage, are insufficient to account for the GMSL rise estimated from tide gauge records<sup>8</sup>, particularly before 1990<sup>7</sup>. For example, a tabulation of contributions to GMSL rise from 1901 to 1990 in the Fifth Assessment Report (AR5; ref. 7) of the Intergovernmental Panel of Climate Change (IPCC) total  $0.5 \pm 0.4 \text{ mm yr}^{-1}$  (90% confidence interval, CI) less than a recent tide gauge derived rate of  $1.5 \pm 0.2 \text{ mm yr}^{-1}$  (90% CI) estimated by Church and White<sup>4</sup> for the same period (the confidence range for this estimate is taken from AR5; refs 7 and 23). Using IPCC terminology, the latter suggests that it is ‘extremely likely’ (probability  $P = 95\%$ ) that GMSL rise from 1901 to 1990 was greater than  $1.3 \text{ mm yr}^{-1}$ , although the bottom-up sum of contributions is ‘likely’ ( $P > 67\%$ ) below this level. The above discrepancy has been attributed to underestimation of almost all possible sources: thermal expansion, glacier mass balance, and Greenland or Antarctic ice sheet mass balance<sup>7,8</sup>.

In this Letter, we revisit the analysis of GMSL since the start of the twentieth century using Kalman smoothing<sup>9</sup> (KS; see Methods). This statistical technique naturally accommodates spatially sparse and temporally incomplete sampling of a global sea-level field, provides a rigorous, probabilistic framework for uncertainty propagation, and can correct for a distribution of GIA and ocean models. We applied the approach to analyse annual records from 622 tide gauges included in the Permanent Service for Mean Sea Level (PSMSL) Revised Local Reference database<sup>15,24</sup> and reconstruct the global field of sea-level change for each year from 1900 to 2010.

To examine the skill with which the KS reconstruction reproduces the tide gauge observations, we compute the time series of residuals at each tide gauge site and examine the distribution of the mean residual (that is, bias) for each site (Fig. 1a). The mean of the mean residuals across all 622 observations is 0.3 mm, with a standard deviation of 5.1 mm, indicating minimal systemic bias.

Comparing reconstructions and tide gauge observations at a selection of individual sites (Fig. 1b–f) shows generally excellent agreement, although there are a small number of outliers. An example outlier is the Champlain tide gauge (Fig. 1f), which has a mean residual of 52 mm. This particular misfit (also evident at other sites in the vicinity) can be attributed to the St Lawrence being a regulated water system where flow is dominated by anthropogenic control rather than global-scale climate dynamics<sup>25</sup>. The eight sites that have mean residuals greater than  $\pm 3\sigma$  (15 mm) from the mean exhibit an average interannual sea-level variability (estimated as the standard deviation after detrending the tide gauge observations) of  $\pm 130 \text{ mm}$ , more than triple the mean inter-annual variability of  $\pm 40 \text{ mm}$  across all sites. Although these outliers

<sup>1</sup>Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Earth and Planetary Sciences, Rutgers University, Piscataway, New Jersey 08854, USA. <sup>3</sup>Rutgers Energy Institute, Rutgers University, New Brunswick, New Jersey 08901, USA.

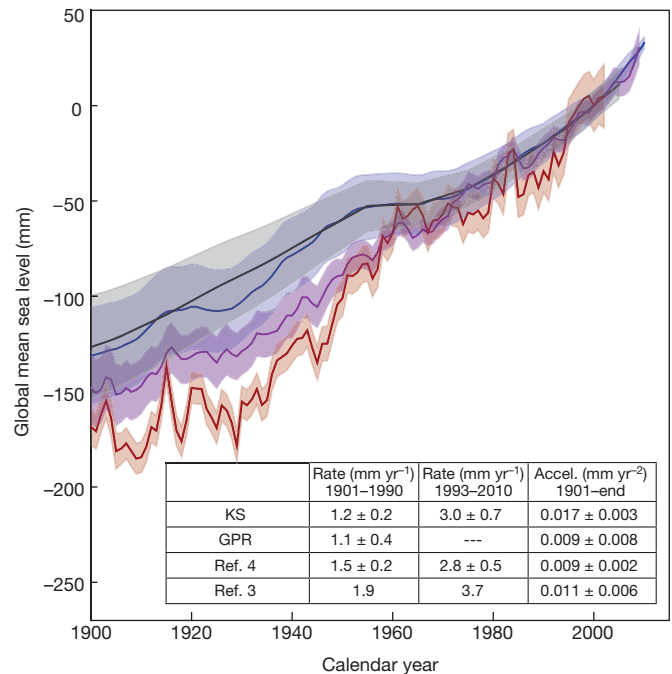


**Figure 1 | Fit of the KS-based reconstruction of sea level to the tide gauge record.** **a**, Histogram of mean residuals (mm) between the sea-level reconstruction and the tide gauge observations at all 622 sites. The mean of all mean residuals is  $0.3 \pm 5.1$  mm ( $\pm 1$  s.d.). **b–f**, Time series of reconstructed annual sea level (black lines, KS mean estimate; grey shading,  $1\sigma$  uncertainty) at New York, USA (**b**), Fremantle, Australia (**c**), Zemlia Bunge, Russia (**d**), Vaasa, Finland (**e**), and Champlain, Canada (**f**), together with the associated annual mean tide gauge observations (red lines).

have large inter-annual variability, the site-specific variability is incorporated into the covariances computed in the probabilistic reconstruction, and the uncertainties in the estimated sea-level trends at these sites reflect this.

The sum of the KS-estimated GMSL changes associated with the mass balance of the Greenland and Antarctic ice sheets, the mass balance of 18 mountain glacier regions, and thermal expansion (Fig. 2, blue line and shading; see Methods) is characterized by an average GMSL rate of  $1.2 \pm 0.2$  mm yr<sup>-1</sup> (90% CI) for 1901–90. As shown in Fig. 3, this is significantly lower than the estimates of  $1.5 \pm 0.2$  mm yr<sup>-1</sup> from Church and White<sup>4</sup> (magenta line in Fig. 2) and  $1.9$  mm yr<sup>-1</sup> from Jevrejeva *et al.*<sup>3</sup> (red line in Fig. 2). The KS-estimated acceleration is  $0.017 \pm 0.003$  mm yr<sup>-2</sup>, larger than our estimates based on the Church and White<sup>4</sup> ( $0.009 \pm 0.002$  mm yr<sup>-2</sup>) and Jevrejeva *et al.*<sup>3</sup> ( $0.011 \pm 0.006$  mm yr<sup>-2</sup>) time series (see Methods).

Church and White<sup>4</sup> combined stationary empirical orthogonal functions (EOFs), computed from  $\sim 20$  years of satellite altimetry data spanning latitudes up to about  $\pm 60^\circ$ , with amplitudes estimated from sparse tide gauge observations. Given the relatively short duration of the altimeter record, the EOFs may be dominated by patterns due to interannual variability rather than the geometry associated with long-term sea-level change<sup>26,27</sup>. Jevrejeva *et al.*<sup>3</sup> used tide gauge records



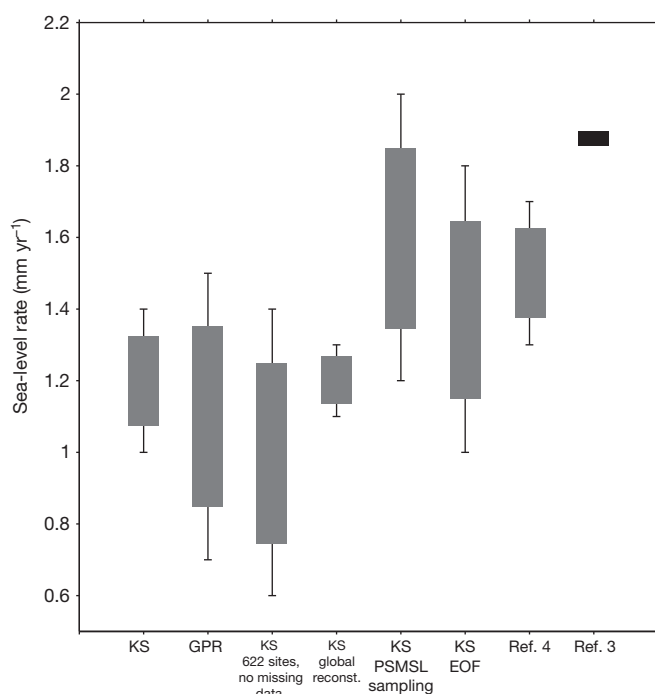
**Figure 2 | Time series of GMSL for the period 1900–2010.** Shown are estimates of GMSL based on KS (blue line), GPR (black line), Church and White<sup>4</sup> (magenta line) and Jevrejeva *et al.*<sup>3</sup> (red line). Shaded regions show  $\pm 1\sigma$  pointwise uncertainty. Inset, trends for 1901–90 and 1993–2010, and accelerations, all with 90% CI. Confidence intervals for Church and White<sup>4</sup> are from refs 7 and 23. Confidence intervals were not available for Jevrejeva *et al.*<sup>3</sup>; data in this reference ends in 2002, so the rate quoted here for 1993–2010 is actually for 1993–2002. Since the GPR methodology outputs decadal sea level, no trend is estimated for 1993–2010. Accelerations are consistently estimated from the KS, GPR, and GMSL time series in refs 3 and 4 (see Methods) from 1901 to the end of each reconstruction.

to compute regional sea-level means and from these computed a global average. Both methodologies involve spatially sparse, temporally incomplete sampling of the global sea-level field, which introduces a potentially significant bias into estimates of GMSL. The KS technique differs from these approaches by using the spatial information inherent in the observations to infer the weights associated with the individual, underlying contributions to the sea-level change. The method extracts global information from the sparse field by taking advantage of the physics-based and model-derived geometry of the contributing processes, thereby reducing the potential for sampling bias.

To understand the origin of the differences between the KS estimate and the higher values of refs 3 and 4, and in particular to quantify the impact of regional binning, spatial sparsity and missing data, we performed several tests.

First, we applied to the KS global sea-level reconstruction a regional binning algorithm similar to that of Jevrejeva *et al.*<sup>3</sup>. In particular, we sampled the reconstruction at the locations of the 622 tide gauge sites, imposed sections of missing data consistent with the PSMSL data availability<sup>15</sup>, binned the tide gauges into 12 ocean regions, and averaged across these regions to compute a GMSL curve. The resulting estimate of the mean GMSL rate from 1901 to 1990 (Fig. 3; 'KS PSMSL sampling'),  $1.6 \pm 0.4$  mm yr<sup>-1</sup> (90% CI), is significantly closer to the estimate of Jevrejeva *et al.*<sup>3</sup>, indicating that combined spatial sparsity and missing data generate an upward bias in estimates of GMSL rates (Fig. 3). Second, we performed a bootstrapping test that repeated the above algorithm for tide gauge subsets ranging from 25 to 600 sites that confirmed this result (see Methods and Extended Data Fig. 3). We also implemented a test to estimate the possible bias in the estimate of GMSL rate introduced in the EOF analysis of Church and White<sup>4</sup> (see Methods; Fig. 3; 'KS EOF'); the result was consistent with the difference between the KS and Church and White<sup>4</sup> results in Fig. 2.



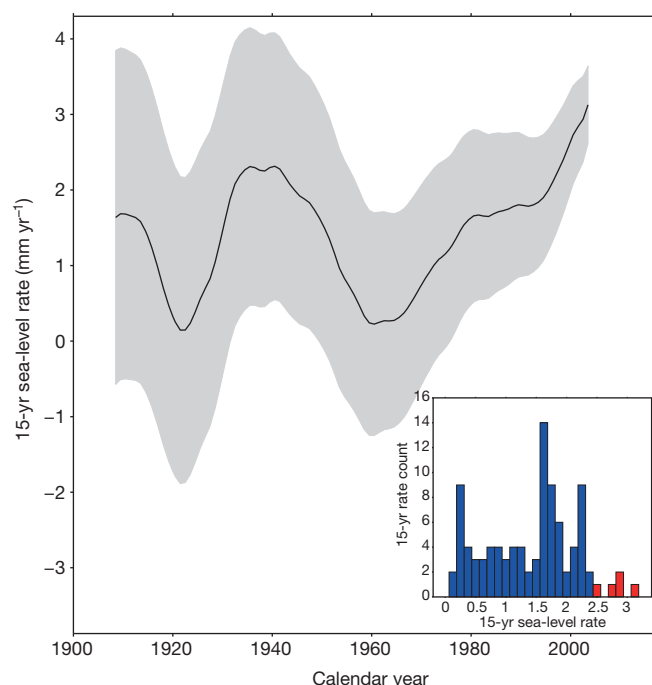


**Figure 3 | Comparison of mean GMSL rates for 1901–90.** Shown are estimates of GMSL rise for the period 1901–90 obtained from six different sampling methods along with previously published rates (see main text for description of each). The box covers the  $1\sigma$  uncertainty range, while the bars represent the 90% CI. In the case of the Jevrejeva *et al.*<sup>3</sup> estimate, uncertainties and confidence intervals were not available.

We performed several other tests to explore the impact of sparsity and missing data on the estimates. Specifically, we applied the binning algorithm as described above but without imposing sections of missing data. The resulting mean GMSL rate estimate for 1901–90 was  $1.0 \pm 0.4 \text{ mm yr}^{-1}$ , close to the KS result (Fig. 3; 'KS 622 sites, no missing data'). Third, we sampled the full reconstruction at a large number of globally distributed sites—that is, the sampling was not confined to the tide gauge sites and no sections of missing data were imposed on the time series—and performed the same regional binning and averaging ('KS global reconstruction'). The resulting rate estimate,  $1.2 \pm 0.1 \text{ mm yr}^{-1}$ , was identical to the KS result (Fig. 3). This indicates that regional binning of estimates, in the absence of sparsity and missing data, does not introduce a significant bias.

To assess the robustness of our probabilistic reanalysis, we also performed a second, independent statistical analysis based on Gaussian process regression<sup>28</sup> (GPR), a technique that also naturally accommodates data sparsity and gaps, and incorporates a suite of GIA and ocean models (see Methods; black line in Fig. 2). The mean GMSL rate for 1901–90 estimated from the GPR analysis,  $1.1 \pm 0.4 \text{ mm yr}^{-1}$ , is consistent with the results of the KS analysis (Fig. 3).

Previous analyses appear to have overestimated the mean GMSL rate over the twentieth century. The KS estimate for the period 1901–90 indicates that it is 'very likely' (probability  $P = 90\%$ ) that the rate of GMSL rise during this period was between 1.0 and  $1.4 \text{ mm yr}^{-1}$ . This estimate closes the sea-level budget for 1901–90 estimated in AR5 (ref. 7) without appealing to an underestimation of individual contributions from ocean thermal expansion, glacier melting, or ice sheet mass balance. Moreover, it may contribute to the ultimate resolution of Munk's sea-level enigma<sup>28</sup> (defined by the argument that Earth rotation measurements and bounds on ocean warming are inconsistent with a rate of sea-level rise beginning in the late nineteenth century of  $1.5\text{--}2.0 \text{ mm yr}^{-1}$ ), since it may lower the signal of twentieth century ice melting in Earth rotation measurements.



**Figure 4 | Moving 15-year averages of GMSL rate estimated using the KS reconstruction of sea level across the entire interval 1901–2010.** The  $x$ -axis represents the mid-point of each 15-year averaging window, and the shading gives the  $1\sigma$  uncertainty range. Inset, histogram of 15-year mean GMSL rate estimates ( $\text{mm yr}^{-1}$ ) for all time windows. The five most recent windows are shown in red.

In contrast, for the period 1993–2010—which coincides with the era of satellite altimetry measurements of sea surface height changes<sup>29</sup>—the KS estimate is consistent with previous results (Fig. 2). The KS estimate,  $3.0 \pm 0.7 \text{ mm yr}^{-1}$  (90% CI), is essentially identical to the tide gauge analysis of Church and White<sup>4</sup> ( $2.8 \pm 0.5 \text{ mm yr}^{-1}$ ; ref. 23). It is also consistent with the estimate based on TOPEX and Jason altimeter measurements ( $3.2 \pm 0.4 \text{ mm yr}^{-1}$ ; ref. 29 as cited by ref. 23 for the period 1993–2010, see also ref. 7).

To assess the anomalous nature of recent sea-level change, we compute 15-year rates through the KS-derived GMSL time series in Fig. 2 from 1901 to 2010. Figure 4 shows both the time series and distribution of these 96 rates, where the 5 most recent time windows are shown in red. The former is in qualitative agreement with a previous inference of multi-decadal trends in acceleration during the twentieth century<sup>30</sup>. While the rates show significant variability, the rate for the 1996–2010 time window,  $3.1 \text{ mm yr}^{-1}$ , is the largest of all computed 15-year rates.

We have revisited twentieth century GMSL rise using probabilistic techniques that combine sea-level records with physics-based and model-derived geometries of the contributing processes. Our estimated GMSL trend for the period 1901–90 ( $1.2 \pm 0.2 \text{ mm yr}^{-1}$ ) is lower than previous estimates, indicating that the rate of GMSL rise during the last two decades represents a more significant increase than previously recognized. Projections of future sea-level rise based on the time series of historical GMSL, notably semi-empirical approaches<sup>11</sup>, should accordingly be revisited.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 July; accepted 14 November 2014.

Published online 14 January 2015.

1. Douglas, B. C. Global sea rise: a redetermination. *Surv. Geophys.* **18**, 279–292 (1997).
2. Holgate, S. J. On the decadal rates of sea level change during the twentieth century. *Geophys. Res. Lett.* **34**, L01602 (2007).

3. Jevrejeva, S., Moore, J. C., Grinsted, A. & Woodworth, P. L. Recent global sea level acceleration started 200 years ago? *Geophys. Res. Lett.* **35**, L08715 (2008).
4. Church, J. A. & White, N. J. Sea level rise from the late 19<sup>th</sup> to the early 21<sup>st</sup> century. *Surv. Geophys.* **32**, 585–602 (2011).
5. Ray, R. D. & Douglas, B. C. Experiments in reconstructing twentieth-century sea levels. *Prog. Oceanogr.* **91**, 496–515 (2011).
6. Wenzel, M. & Schroeter, J. Reconstruction of regional mean sea level anomalies from tide gauges using neural networks. *J. Geophys. Res.* **115**, C08013 (2010).
7. Church, J. A. *et al.* in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) Ch. 13 (Cambridge Univ. Press, 2013).
8. Gregory, J. M. *et al.* Twentieth-century global mean sea level rise: is the whole greater than the sum of the parts? *J. Clim.* **26**, 4476–4499 (2013).
9. Hay, C. C., Morrow, E., Kopp, R. E. & Mitrovica, J. X. Estimating the sources of global sea level rise with data assimilation techniques. *Proc. Natl Acad. Sci. USA* **110**, 3692–3699 (2013).
10. Kopp, R. E. Does the mid-Atlantic United States sea level acceleration hot spot reflect ocean dynamic variability? *Geophys. Res. Lett.* **40**, 3981–3985 (2013).
11. Rahmstorf, S. A semi-empirical approach to projecting future sea-level rise. *Science* **315**, 368–370 (2007).
12. Ekman, M. The world's longest continued series of sea level observations. *Pure Appl. Geophys.* **127**, 73–77 (1988).
13. Woodworth, P. L. High waters at Liverpool since 1768: the UK's longest sea level record. *Geophys. Res. Lett.* **26**, 1589–1592 (1999).
14. Wöppelmann, G., Pouvreau, N., Coulomb, A., Simon, B. & Woodworth, P. L. Tide gauge datum continuity at Brest since 1711: France's longest sea-level record. *Geophys. Res. Lett.* **35**, <http://dx.doi.org/10.1029/2008GL035783> (2008).
15. Holgate, S. J. *et al.* New data systems and products at the permanent service for mean sea level. *J. Coast. Res.* **29**, 493–504 (2013).
16. Peltier, W. R. & Tushingham, A. M. Global sea level rise and the greenhouse effect: might they be connected? *Science* **244**, 806–810 (1989).
17. Davis, J. L. & Mitrovica, J. X. Glacial isostatic adjustment and the anomalous tide gauge record of eastern North America. *Nature* **379**, 331–333 (1996).
18. Woodward, R. On the form of position of mean sea level. *US Geol. Surv. Bull.* **48**, 87–170 (1888).
19. Clark, J. A. & Lingle, C. S. Future sea level changes due to West Antarctic ice sheet fluctuations. *Nature* **269**, 206–209 (1977).
20. Mitrovica, J. X., Tamisiea, M. E., Davis, J. L. & Milne, G. A. Recent mass balance of polar ice sheets inferred from patterns of global sea-level change. *Nature* **409**, 1026–1029 (2001).
21. Conrad, C. & Hager, B. H. Spatial variations in the rate of sea level rise caused by present day melting of glaciers and ice sheets. *Geophys. Res. Lett.* **24**, 1503–1506 (1997).
22. Kopp, R. E. *et al.* The impact of Greenland melt on local sea levels: a partially coupled analysis of dynamic and static equilibrium effects in idealized water hosing experiments. *Clim. Change* **103**, 619–625 (2010).
23. Rhein, M. *et al.* in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) Ch. 3 (Cambridge Univ. Press, 2013).
24. Permanent Service for Mean Sea Level (PSMSL). Tide gauge data. <http://www.psmsl.org/data/obtaining/> (2014); retrieved 31 March 2014.
25. Talbot, A. (ed.) *Water Availability Issues for the St. Lawrence River: An Environmental Synthesis* (Environment Canada, Montreal, 2006).
26. Calafat, F. M., Chambers, D. P. & Tsimplis, M. N. On the ability of global sea level reconstructions to determine trends and variability. *J. Geophys. Res.* **119**, 1572–1592 (2014).
27. Christiansen, B., Schmith, T. & Thejll, P. A surrogate ensemble study of sea level reconstructions. *J. Clim.* **23**, 4306–4326 (2010).
28. Munk, W. Twentieth century sea level: an enigma. *Proc. Natl Acad. Sci. USA* **99**, 6550–6555 (2002).
29. Nerem, R. S., Chambers, D. P., Choe, C. & Mitchum, G. T. Estimating mean sea level from the TOPEX and Jason altimeter missions. *Mar. Geod.* **33**, 435–446 (2010).
30. Woodworth, P. L. *et al.* Evidence for the accelerations of sea level on multi-decadal and century timescales. *Int. J. Climatol.* **29**, 777–789 (2009).

**Acknowledgements** Tide gauge data were provided by PMSL ([www.psmsl.org](http://www.psmsl.org)). This work was supported by US National Science Foundation grants ARC-1203414 and ARC-1203415, the New Jersey Sea Grant Consortium and the National Oceanic and Atmospheric Administration (NJS GC project 6410-0012), Rutgers University (R.E.K., C.C.H.), and Harvard University (J.X.M., C.C.H. and E.M.). We thank P. Woodworth for comments on earlier versions of this manuscript.

**Author Contributions** C.C.H. and E.M. developed the methodology and performed the analysis. R.E.K. and J.X.M. helped in the study design. All authors contributed to the discussion and writing of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.C.H. ([carlinghay@fas.harvard.edu](mailto:carlinghay@fas.harvard.edu)).

## METHODS

**Probabilistic estimation methods.** Kalman smoothing (KS) and Gaussian process regression (GPR), both discussed in detail below, share three advantages over the approaches taken in traditional tide gauge analyses. First, the Bayesian nature of both approaches naturally accommodates the spatiotemporal changes in the availability of the sea-level records (that is, sparsity and missing data). Second, the probabilistic approaches correct for a distribution of GIA and ocean models rather than adopting a specific model for each process, and they thus reduce a potentially important bias in previous estimates of the GMSL change<sup>17,31</sup>. Last, as both methods are fully probabilistic, they allow for the propagation of measurement and inferential uncertainties and correlations throughout the complete analysis time period. Despite these commonalities, the implementations of KS and GPR differ significantly.

**Kalman smoother.** The KS methodology is divided into four steps<sup>9</sup>, the first three of which are repeated by employing the spatial fields of GIA and ocean dynamic models from all possible combinations of 161 different Earth rheological models and 6 global climate model (GCM) simulations from CMIP5 (ref. 32) (see below for details of the rheological and climate models). First, a priori model estimates of both local sea level and the individual mass contributions from the Greenland, West Antarctic and East Antarctic ice sheets, as well as 18 major mountain glacier regions, are recursively corrected by tide gauge observations as the estimates are propagated forward through time. The local sea level is linked to the individual mass contributions through the unique spatial patterns, or ‘fingerprints,’ of sea-level change associated with rapid mass loss from land-based ice<sup>18–21</sup>. The forward step yields an estimate of local sea level and land ice contributions at each time slice, conditional on all earlier observations and a particular combination of GIA and GCM models. Second, the procedure is run backward in time, with the initial state estimate being the last estimate from the first step. The third, smoothing step optimally combines the results of the first two passes based upon the uncertainties of the respective estimates. The result is an estimate of local sea levels and land ice contributions conditional upon the entire set of observations and specific pairings of GIA and GCM models. Finally, the results from different GIA/GCM combinations are linearly combined, weighted by their likelihood, to yield an a posteriori probability distribution for local sea levels and land ice contributions, conditional upon the tide gauge observations.

A comprehensive discussion of our application of the KS technique to the analysis of tide gauge measurements is given in ref. 9, which also includes synthetic tests to assess the performance of the procedure. Several subsequent refinements of this approach are summarized below.

Reference 9 defined the state vector to include estimates of sea level at every tide gauge site, the mass loss rates of three ice sheets, and the temporally correlated noise in the sea-level observations. Using only tide gauge observations limits our ability to separate estimates of sea level from estimates of the temporally correlated noise. This led us to modify the KS approach in two ways.

First, the state vector includes only an estimate of total sea level at every tide gauge site in addition to the desired mass loss rates. This yields the following state vector,  $\mathbf{x}_k$ , at every time step,  $k$ :

$$\mathbf{x}_k = [\mathbf{h}_k \quad \mathbf{B}_k]^T$$

where  $\mathbf{h}_k$  is a vector of sea level at the 622 tide gauge sites, and  $\mathbf{B}_k$  is a vector containing the scalar weightings of 3 ice sheets and 18 mountain glacier regions (see below), as well as a uniform component that accounts for global mean thermal expansion and any additional mass contributions from smaller mountain glaciers.

Second, while in ref. 9 the observation model consisted of the sum of the estimated sea level, correlated noise, and white noise, here, the observation model consists only of the estimated sea level and white noise at each tide gauge site. Temporal correlations due to ocean dynamics are now modelled by the annual, spatial, CMIP5 ocean model fields (see below for a more detailed description of the CMIP5 model fields).

Sea level is modelled as the Euler integration of the contributions from melt sources,  $\mathbf{B}_{k-1}\mathbf{y}$ , (with  $\mathbf{y}$  being the matrix of sea-level fingerprints associated with rapid land-ice mass loss), the ongoing rate of sea-level change due to GIA,  $\mathbf{G}$ , and the rate of change of sea level due to ocean dynamics,  $\dot{\mathbf{S}}_{k-1}$ , from the spatial fields in the CMIP5 model outputs:

$$\mathbf{h}_k = \mathbf{h}_{k-1} + \Delta t (\mathbf{B}_{k-1}\mathbf{y} + \mathbf{G} + \dot{\mathbf{S}}_{k-1}) + \mathbf{w}_h$$

where  $\mathbf{w}_h$  represents a zero-mean, white noise term associated with sea level.

The scalar weightings of the fingerprints are modelled as a random walk:

$$\mathbf{B}_k = \mathbf{B}_{k-1} + \mathbf{w}_B$$

where  $\mathbf{w}_B$  represents a zero-mean, white noise term associated with the melt contributions. The forward filtering pass of the Kalman smoother follows the steps

outlined in ref. 9. A final departure from the methodology presented in ref. 9 is that we implemented a three-pass fixed-interval smoother<sup>33</sup> in place of a Rauch-Tung-Stein two-pass smoother<sup>34</sup>.

**Gaussian process regression.** The GPR approach, in contrast, models sea level as a multivariate Gaussian field defined by spatiotemporal mean and covariance functions that describe the underlying processes responsible for sea-level variability. Specifically, Gaussian process priors describing the contributions from land ice, GIA, and ocean models are conditioned simultaneously upon the available observations to produce the conditional, posterior distribution of sea level at decadal intervals throughout the twentieth century. In contrast to the KS, the GPR approach directly estimates the intertemporal covariance of the posterior; the associated computational demands require the use of decadal rather than annual means. Rather than being based upon discrete GIA and GCM models as in the KS approach, the GPR approach employs Gaussian process priors for the GIA and ocean dynamics contributions that are estimated, respectively, from the 161 GIA model predictions and 6 GCM outputs (see below). The distribution describing each land ice mass contribution is modelled assuming a prior spatio-temporal covariance, with the temporal component estimated from previous, non-sea level based estimates of land ice melt and the spatial component from the sea-level fingerprints associated with the melt source.

We model decadal-average sea level as a spatiotemporal field:

$$f(\mathbf{x}, t) = f_{\text{GIA}}(\mathbf{x}, t) + f_{\text{M}}(\mathbf{x}, t) + f_{\text{LSL}}(\mathbf{x}, t)$$

where  $f_{\text{GIA}}$ ,  $f_{\text{M}}$ , and  $f_{\text{LSL}}$  are respectively the components of sea level due to ongoing GIA, land ice mass loss, and local effects associated with ocean dynamics, tectonics and other non-climatic factors, each as a function of location,  $\mathbf{x}$ , and time,  $t$ . Each sea-level component is modelled as a Gaussian process with a prior mean function,  $\mu_i(\mathbf{x}, t)$ , and covariance function,  $\mathbf{K}_i(\mathbf{x}, t, \mathbf{x}', t')$ .

The total field can be partitioned into observed sites,  $\mathbf{f}_1$ , and unobserved sites,  $\mathbf{f}_2$ , and subsequently written as a joint, multivariate distribution, such that:

$$\begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{12}^T & \mathbf{K}_{22} \end{bmatrix} \right)$$

Observations,  $\mathbf{y}$ , are modelled as the underlying sea-level field with additive white noise characterized by zero mean and a covariance  $\Sigma_p$ , such that the joint distribution becomes:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_2 \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{11} + \Sigma_p & \mathbf{K}_{12} \\ \mathbf{K}_{12}^T & \mathbf{K}_{22} \end{bmatrix} \right)$$

Using standard statistical results (see, for example, ref. 35), the posterior mean and covariance,  $\mathbf{f}_2$  and  $\mathbf{V}_2$ , of the unobserved field conditioned upon the observations are:

$$\mathbf{f}_2 = \mu_2 + \mathbf{K}_{12}^T [\mathbf{K}_{11} + \Sigma_p]^{-1} \mathbf{y}$$

and

$$\mathbf{V}_2 = \mathbf{K}_{22} - \mathbf{K}_{12}^T [\mathbf{K}_{11} + \Sigma_p]^{-1} \mathbf{K}_{12}$$

To estimate the underlying constituents of the total sea-level field, the prior mean and covariance of the unobserved field (that is,  $\mu_2$ ,  $\mathbf{K}_{12}$ ,  $\mathbf{K}_{22}$ ) are set to the distribution of the desired quantity alone. For example, setting  $\mu_2$ ,  $\mathbf{K}_{12}$ , and  $\mathbf{K}_{22}$  equal to  $\mu_2^{\text{M}}$ ,  $\mathbf{K}_{12}^{\text{M}}$ ,  $\mathbf{K}_{22}^{\text{M}}$ , returns the posterior mean and covariance of sea-level change due to the melt contributions. Once all the underlying constituent sea-level fields are determined, the global mean of those components can be computed and added to estimate GMSL.

The elements of the prior covariance matrix of the melt contribution,  $\mathbf{K}^{\text{M}}$ , are defined as:

$$K_{ij}^{\text{M}} = \sum_{a=1}^n \left( A_{i,j,a}^{\text{M,L}} + A_{i,j,a}^{\text{M,RQ}} \right) \left( B_{i,j,a}^{\text{M}} \right)$$

where the subscripts indicate the  $i$ th row and  $j$ th column element of the  $a$ th ice sheet or mountain glacier. The time dependence of the covariance matrix is taken to be the sum of a linear component,  $A^{\text{M,L}}$ , which accounts for secular changes in the melt contributions, and a rational quadratic term,  $A^{\text{M,RQ}}$ , that represents a smoothly-varying function of variability:

$$A^{\text{M,L}}(t_q, t_p) = k_1 t_q t_p$$

and



$$A^{M,RQ}(t_q, t_p) = k_2 \left( 1 + \frac{\Delta t_{q,p}^2}{2\alpha\tau_s^2} \right)^{-\alpha}$$

Here,  $t_q$  and  $t_p$  represent the time at the  $q$ th and  $p$ th time step,  $\Delta t_{q,p}$  represents the time difference between these steps, and  $k_1$ ,  $k_2$ ,  $\alpha$ , and  $\tau_s$  are hyperparameters that define the linear amplitude, rational quadratic amplitude, roughness, and characteristic timescale of the covariance functions<sup>35</sup>. To estimate the hyperparameters we adopt an empirical Bayesian approach where we compute the parameters that maximize the likelihood of reconstructed time series of previous mountain glacier estimates<sup>36</sup> and ice sheet estimates<sup>37</sup>.

The spatial weighting of the prior covariance,  $B^M$ , is computed as the outer product of the unique fingerprint associated with melt from the corresponding land-based ice source.

The prior spatiotemporal mean and covariance for the GIA contribution to sea-level change,  $\mu^{GIA}(\mathbf{x}, t)$  and  $K^{GIA}(\mathbf{x}, t)$ , respectively, are taken as the sample mean and covariance of the 161 predictions of sea-level change described below.

The distribution of the contribution to sea-level changes from thermosteric and ocean dynamic effects is partially modelled as the sample mean and covariance of the CMIP5 model outputs<sup>32</sup>. However, since a small number of models are used to compute the distribution statistics, the estimated distribution may not be representative of the parent distribution. Consequently, we augment the sample covariance with a space-time separable covariance structure consisting of the product of two Matérn functions<sup>35</sup>,  $C$ : one representing the temporal distribution and the other representing the spatial, such that the total prior covariance describing local sea-level change is given by:

$$K^{LSL}(\mathbf{x}, t) = K^{CMIP5}(\mathbf{x}, t) + C(t, v_1, \tau) C(\mathbf{x}, v_2, L)$$

where  $K^{CMIP5}$  is the sample covariance of the CMIP5 model outputs,  $v_1$  and  $\tau$  are the smoothness parameter and characteristic timescale of the temporal Matérn function, respectively, and  $v_2$  and  $L$  are the smoothness parameter and characteristic length scale of the spatial Matérn function, respectively. For the exponents within the Matérn functions we follow ref. 10 and set the exponent on the spatial component to  $v_2 = 5/2$  (reflecting a relatively smooth, twice-differentiable field) and the exponent on the temporal component to  $v_1 = 3/2$  (reflecting a once-differentiable time series, in which rate is always defined but can change abruptly). As with the melt covariance hyperparameters, we use an empirical Bayesian approach to estimate the maximum-likelihood time and length scales of the Matérn functions to be 46 years and 90 km, respectively. Note that there is some trade-off between the Matérn exponent values and the hyperparameter characteristic scales: the selection of, say, a lower exponent (giving rise to a less smooth functional form) would result in a longer length scale.

In addition to capturing the inaccuracies of the ocean dynamics distribution, the Matérn functions also model local tectonic, geomorphological and other non-climatic contributions to local sea-level change. These hyperparameters, and a white-noise variance, are computed by finding the parameters that maximize the likelihood of the available tide gauge observations given the complete sea-level model.

**Sea-level fingerprints.** Extended Data Fig. 1a and b shows global maps of sea-level change, known as sea-level fingerprints, associated with rapid, uniform mass loss across the Greenland Ice Sheet (GIS) and the West Antarctic Ice Sheet (WAIS), respectively. The sea-level changes are normalized by the equivalent GMSL change. Both fingerprints are characterized by a large amplitude sea-level fall in the region adjacent to the melting ice sheet with a gradual rise in sea level moving away from the ice sheet. The computation of the fingerprints is based upon a gravitationally self-consistent sea-level theory that takes into account shoreline migration and changes in grounded, marine-based ice cover as well as the impact on sea level of perturbations in the Earth's rotation axis<sup>38–40</sup>.

In addition to the GIS and WAIS, fingerprints were computed for the East Antarctic Ice Sheet (EAIS) and glaciers of Alaska, the Alps, Baffin Island, the Caucasus, Ellesmere Island, Franz Josef Land, High Mountain Asia, Altai, Iceland, Kamchatka, the low-latitude Andes, New Zealand, Novaya Zemlya, Patagonia, Scandinavia, Severnaya Zemlya, Svalbard, and Western Canada/US.

We also include a spatially uniform pattern to account for changes in GMSL due to land ice sources not included in the above set of glaciers. In the Kalman smoother, this uniform 'fingerprint' also captures changes in GMSL due to globally uniform thermal expansion and terrestrial water storage variations<sup>9</sup>.

**GIA models.** The first step when analysing tide gauge records is to correct for sea-level contributions due to the ongoing GIA of the Earth in response to the ice age cycles. Predictions of GIA are dependent on the geometry and deglaciation history of the Late Pleistocene ice sheets and the Earth's viscoelastic structure. In this study, we computed 160 different GIA predictions distinguished on the basis of the adopted lower-mantle viscosity, upper-mantle viscosity, and thickness of a high-viscosity (effectively elastic) lithosphere. Additionally, we computed a GIA prediction using

the VM2 viscosity profile<sup>41</sup>. These were combined with the ICE-5G (Ref. 41) global ice sheet reconstruction for the last glacial cycle. A detailed description of physical processes that contribute to the total GIA signal can be found in ref. 42.

We adopted values for the three rheological model parameters that encompass all recent estimates of the Earth's structure. The lower-mantle viscosity was varied in the range  $(2-100) \times 10^{21}$  Pa s, upper-mantle viscosity in the range  $(0.3-1) \times 10^{21}$  Pa s, and lithospheric thickness in the range 72–150 km. Extended Data Fig. 2a and b shows the mean and standard deviation of the model predictions. The largest variance is seen in the region within the near field of the former ice sheets, including areas of ancient ice cover and the so-called peripheral bulges.

**Ocean dynamics models.** We treat the thermosteric and ocean dynamic contributions to sea level using the historical experiment output from 6 global climate models of the World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project phase 5 (CMIP5) data set<sup>32</sup>. Following ref. 9, the models we use are: bcc-csm1-1 from the Beijing Climate Center, CanESM2 from Environment Canada, the NOAA-GFDL model GFDL-ESM2M, the Institut Pierre Simone Laplace IPSL-CM5A-LR model, MRI-CGCM3 from the Japanese Meteorological Institute, and NorESM1-M from the Norwegian Climate Centre. For the KS methodology, we use the zero-mean spatial field 'zos' that is supplied by all the models. In the GPR, we add to 'zos' each model's estimated globally averaged sea-level change due to thermal expansion: 'zossga'.

While the CMIP5 model outputs are provided as global ocean grids, the field values at the specific locations of tide gauges are required, as input, to both the KS and GPR analyses. Where the tide gauges are coincident with model grid points, the associated value of the model output is used. Otherwise, an inverse distance weighting interpolation scheme is used to estimate the field at the desired location.

We examined three alternative interpolation schemes to assess the sensitivity of the KS GMSL estimate to this choice: (1) a nearest-neighbour approach, selecting the value on the CMIP5 grid that is closest to the tide gauge site; (2) a Delaunay interpolant, computing a linear interpolation between the irregularly spaced model cells along the coastlines; and (3) a Gaussian process (or simple kriging) methodology. For the Gaussian process interpolation, we employed a Gaussian process prior with a mean equal to the mean of the model grid values within a 200 km radius of the tide gauge location and a Matérn covariance function with smoothness parameter equal to 5/2. Since we are interested in the variability of the ocean models immediately surrounding each tide gauge site, the length scale of the Matérn covariance function was set to  $1^\circ$  ( $\sim 110$  km). Neither the nearest-neighbour approach nor the Delaunay interpolated altered the estimate of the GMSL rate over the time period 1901–90. The Gaussian process interpolation scheme changed the GMSL estimate by less than 2%, significantly smaller than the estimated  $\pm 0.2 \text{ mm yr}^{-1}$  90% CI on the estimate.

**Computation of GMSL rates and accelerations.** The mean and uncertainty of GMSL rates are estimated using a generalized least squares regression of a linear trend to the reconstructed GMSL time series. While the GPR methodology outputs a full temporal covariance matrix, the KS methodology does not. For this purpose, we adopt a temporal covariance matrix  $\mathcal{Z}$  with elements having the form:

$$\Sigma_{i,j} = \sigma_i \sigma_j e^{-\frac{(t_j - t_i)}{\tau}}$$

where  $\sigma_i$  and  $\sigma_j$  are the instantaneous uncertainties in GMSL at time  $i$  and  $j$ , respectively, derived in the multi-model KS analysis. To estimate the decorrelation timescale,  $\tau$ , we examined the annual PSMSL tide gauge data and computed the mean temporal correlation coefficient across all tide gauges. This coefficient approaches zero after 2 years, and we set  $\tau$  to 3 years. Estimates of acceleration in GMSL cited in the main text for the two probabilistic analyses are computed using a generalized least squares fit of a quadratic through the associated GMSL time series. Estimates of acceleration for the Church and White<sup>4</sup> and Jevrejeva *et al.*<sup>3</sup> time series listed in Fig. 2 are based on a weighted least squares regression through the published time series (see figure legend).

**Analysis of bias introduced by using a subset of tide gauges.** We used a bootstrapping technique to assess the potential biases introduced in estimates of GMSL rates when only a subset of tide gauge records is used. We randomly sampled our global sea-level reconstruction based on the Kalman smoother at a specific number of tide gauge sites in the database of 622 sites, computed the associated GMSL curve by binning the sites into 12 regions and averaging the result, and then used this curve to determine the rate of sea-level change over the time period 1901–90. The time series of the sea-level reconstruction at any given tide gauge site were sampled to match any missing data at that site in the PSMSL database. We repeated the analysis 100 times for subsets ranging in size from 25 to 600 sites. The mean sea-level rate we computed in this exercise and its associated uncertainty are shown in Extended Data Fig. 3 as a function of the number of sites. The horizontal blue line and shading is the mean rate of sea-level rise from 1901 to 1990, and its associated

uncertainty, respectively, obtained from the KS-derived time series ( $1.2 \pm 0.2 \text{ mm yr}^{-1}$ ; Figs 2, 3).

The mean sea-level rate obtained from this analysis asymptotes towards its final value and the spread in rates decreases monotonically as the number of tide gauges used in the analysis increases. The asymptote lies  $\sim 0.4 \text{ mm yr}^{-1}$  above the KS estimate, which is consistent with the difference between the KS and 'KS PSMSL sampling' rate estimates for 1901–90 shown in Fig. 3. This result suggests that the combined effects of data sparsity and missing data introduce an upward bias into the estimate of GMSL. This bias is reduced in the KS (and GPR) methodologies because these techniques extract global information by using the observations, together with model-based geometries (or covariances) associated with the underlying contributions, to estimate (and sum) these contributions.

**Analysis of bias introduced by an EOF analysis of altimetry records.** To compare our results with EOF-based reconstructions of sea level<sup>44,3</sup>, we computed the GMSL time series following the approach adopted by Church and White<sup>4</sup>, but replacing altimetry and tide gauge observations with our KS reconstruction. The EOFs were computed using the KS sea-level reconstruction from 1993 to 2010, limited to the latitudinal observation range of satellite altimetry ( $65^\circ \text{ N}$  to  $65^\circ \text{ S}$ ). As in ref. 43, a spatially uniform EOF was added to the basis set to account for changes in mean sea level within the altimetry data (here the KS reconstruction), while the weights of the EOFs were computed using the first differences of the KS reconstruction at the tide gauge locations (sampled to reflect missing data in the PSMSL database) in order to eliminate dependence on a consistent datum. The GMSL time series was computed using an area-weighted mean of the EOF-reconstruction. To compute the uncertainty in our estimated GMSL, we sampled our distribution for each KS reconstructed tide gauge 1,000 times and computed the corresponding EOF-derived GMSL time series. We used this distribution of GMSL curves with a generalized least squares regression to compute a trend and uncertainty. This analysis yielded a linear trend of  $1.4 \pm 0.4 \text{ mm yr}^{-1}$ , demonstrating the existence of a bias since the 'true' underlying reconstruction has a trend of  $1.2 \pm 0.2 \text{ mm yr}^{-1}$  (see Fig. 3, 'KS EOF').

**Inverted barometer correction.** The results in the manuscript were obtained using tide gauge observations that were not corrected for the inverse barometer (IB) effect. Previous studies (for example, refs 44 and 45) have shown that the sea-level response to atmospheric pressure changes can be non-negligible on regional scales.

In order to investigate the potential effect that atmospheric pressure changes have on our probabilistic estimate of GMSL, we repeated the KS analysis on the full tide gauge data set after we corrected these records for the IB effect. Specifically, we used the HadSLP2 global reconstructed atmospheric pressure data set<sup>46</sup> to compute the IB correction. We next applied the correction to the observations at the 622 tide gauge sites and then re-ran the KS analysis. The 1901–90 GMSL rate of change associated with this analysis is  $1.2 \pm 0.2 \text{ mm yr}^{-1}$ , consistent with the value cited in the main text. We conclude that while the IB effect can impact regional sea-level histories, it has a negligible effect on our probabilistic estimates of GMSL.

**Optimality of the Kalman smoother.** Local sea levels observed by tide gauges reveal significant interannual and decadal variability. This variability can lead to temporal correlation in the sea-level time series that needs to be considered if one seeks to obtain optimal estimates of the underlying GMSL contributions. In order to test the optimality of the Kalman smoother, we investigated the properties of the innovation sequence by computing the residuals between the observations and the KS model estimate of sea level at every tide gauge site. Since every KS estimate of sea level is accompanied by its associated uncertainty, we randomly sampled from each sea-level distribution to obtain 100 time series of residuals for every site. Following the optimality test described in ref. 47, we computed the mean AR(1) coefficient across the 100 samples at each tide gauge site. An optimal Kalman smoother is characterized by a white noise innovation sequence. In practice, this means that, within uncertainty, the AR(1) coefficients of the innovation sequences will be close to zero. In the exercise above, we obtained a mean AR(1) coefficient of  $0.2 \pm 0.3$  (90% confidence). This indicates that our innovation sequence is (within uncertainty) white noise and that the smoother is, or is close to, optimal.

**Sensitivity of GMSL estimates to limitations of the CMIP5 climate simulations.** The presence of unmodelled ocean dynamics can also affect the smoother performance. As described above, the limitations of the CMIP5 simulations as models for the true dynamic variability of the oceans is addressed in the GPR analysis by augmenting the covariance computed from the climate runs with two additional terms: a covariance modelled with two Matérn functions, and a white noise variance.

To assess the sensitivity of the KS analysis to unmodelled ocean dynamics, we examined its response to (1) a known synthetic ocean dynamic signal and (2) the inclusion of the dynamic response to freshwater hosing of the North Atlantic.

We used the mean KS estimates of the ice sheet melt rates and uniform sea-level contribution, as well as the multi-model estimate of the GIA contribution, to construct synthetic sea-level observations at the 622 tide gauge sites. We then added the dynamic sea-level change associated with one of the six CMIP5 climate models and ran the multi-model KS using the five remaining climate models to obtain an

estimate of the GMSL rate. We repeated this analysis for each of the CMIP5 simulations. By not including the climate model used in constructing the synthetics in the multi-model component of the KS methodology, we tested the ability of the smoother to account for unmodelled dynamics. The 1901–90 GMSL rates determined from the complete set of 6 analyses ranged from  $1.1$  to  $1.3 \text{ mm yr}^{-1}$ . Five of these analyses yielded a 90% CI of  $0.2 \text{ mm yr}^{-1}$ , while the sixth yielded a 90% CI of  $0.3 \text{ mm yr}^{-1}$ . These values are consistent with the results for the KS analysis cited in the main text ( $1.2 \pm 0.2 \text{ mm yr}^{-1}$ ).

To assess the sensitivity of our GMSL results to ocean dynamic effects due to freshwater input ('hosing') from GIS melt, we used the results of a previous study<sup>48</sup> to investigate the dynamic sea-level signal arising from North Atlantic freshwater 'hosing' simulations. Specifically, we computed the difference between the results of the 0.1 Sv hosing run and the control (no-hosing) simulation described in ref. 48 and scaled this difference by 0.05 to approximate a synthetic dynamic signal for a GIS melt rate equivalent to  $0.5 \text{ mm yr}^{-1}$  GMSL rise over the twentieth century. After subtracting a uniform  $0.5 \text{ mm yr}^{-1}$  from the spatial pattern, we calculated time series of this signal at all 622 tide gauge sites, added these to the observed record, and repeated the KS analysis. The presence of these unmodelled dynamics has negligible effect on our estimate of GMSL. The 1901–90 rate estimated in the above test agrees with the value presented in the manuscript ( $1.2 \pm 0.2 \text{ mm yr}^{-1}$ ).

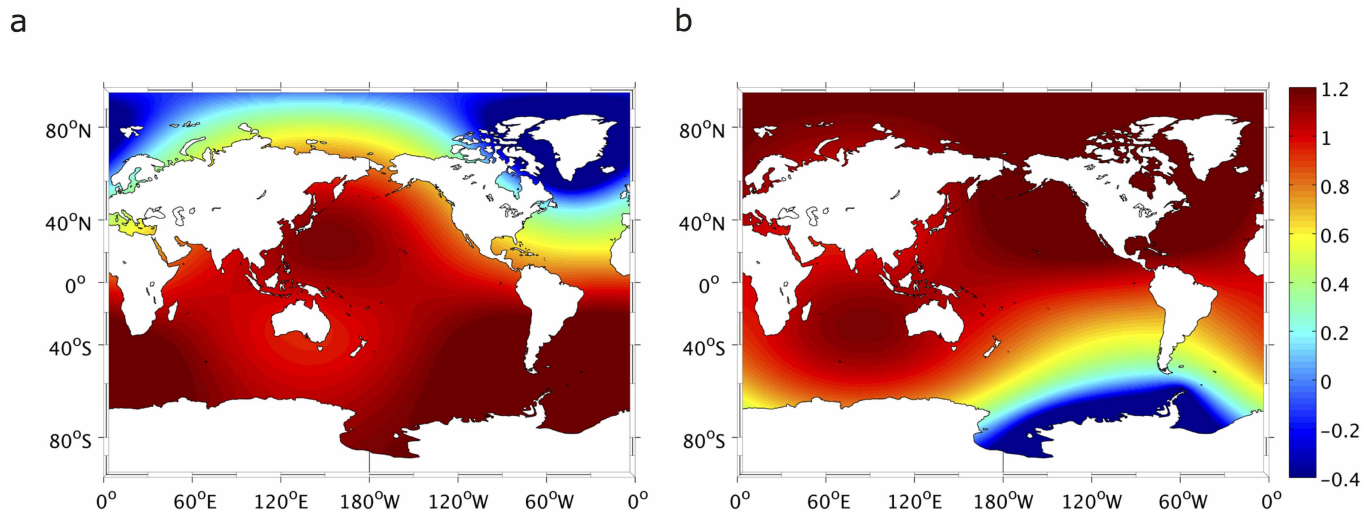
While the above sensitivity tests indicate that the probabilistic analyses have quantified, with reasonable accuracy, the impact of uncertainties in CMIP5 models of ocean dynamic variability, improving such models is an important requirement in any effort to further refine estimates of GMSL rates.

**Kalman smoother reconstruction of sea level at sites with no observations.** To investigate how well the KS is able to estimate sea level at sites without observations, we ran the Kalman smoother using data from 450 randomly chosen tide gauge sites and estimated the sea level at the remaining 172 sites. Extended Data Fig. 4 shows the GMSL time series estimated in this new analysis as well as a comparison of the estimated and observed sea level at a representative subset of 5 of these 172 sites (the remaining sites show similar fits). We calculate a 1901–90 GMSL rate of  $1.2 \pm 0.2 \text{ mm yr}^{-1}$ , consistent with the results presented in the manuscript when all 622 tide gauge sites are used in the analysis. The consistency between the estimated and observed values at the 172 tide gauge sites also indicates that limitations of the CMIP5 simulations in modelling ocean dynamics are not degrading the ability to predict sea-level trends at sites without observations. More generally, the analysis demonstrates the power of the KS method in reconstructing sea level when the method is applied with physics-based and model-derived geometries of the underlying physical processes.

31. Mitrovica, J. X. & Davis, J. L. Present-day post-glacial sea level change far from the Late Pleistocene ice sheets: implications for recent analyses of tide gauge records. *Geophys. Res. Lett.* **22**, 2529–2532 (1995).
32. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
33. Gelb, A., Kasper, J. F., Nash, R. A., Price, C. F. & Sutherland, A. A. in *Applied Optimal Estimation* (ed. Gelb, A.) Ch. 5 (MIT Press, 1974).
34. Rauch, H. E., Tung, F. & Striebel, C. T. Maximum likelihood estimates of linear dynamic systems. *Am. Inst. Aeronaut. Astronaut. J.* **3**, 1445–1450 (1965).
35. Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning* (MIT Press, 2006).
36. Kaser, G., Cogley, J. G., Dyurgerov, M. B., Meier, M. F. & Ohmura, A. Mass balance of glaciers and ice caps: consensus estimates for 1961–2004. *Geophys. Res. Lett.* **33**, L19501 (2006).
37. Gregory, J. M. *et al.* Twentieth-century global mean sea level rise: is the whole greater than the sum of the parts? *J. Clim.* **26**, 4476–4499 (2013).
38. Kendall, R. A., Mitrovica, J. X. & Milne, G. A. On post-glacial sea level – II. Numerical formulation and comparative results on spherically symmetric models. *Geophys. J. Int.* **161**, 679–706 (2005).
39. Milne, G. A. & Mitrovica, J. X. Postglacial sea-level change on a rotating Earth. *Geophys. J. Int.* **133**, 1–19 (1998).
40. Mitrovica, J. X. *et al.* On the robustness of predictions of sea level fingerprints. *Geophys. J. Int.* **187**, 729–742 (2011).
41. Peltier, W. R. Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and GRACE. *Annu. Rev. Earth Planet. Sci.* **32**, 111–149 (2004).
42. Mitrovica, J. X. & Milne, G. A. On the origin of late Holocene sea-level highstands within equatorial ocean basins. *Geophys. J. Int.* **21**, 2179–2190 (2002).
43. Church, J. A., White, N. J., Coleman, R., Lambeck, K. & Mitrovica, J. X. Estimates of the regional distribution of sea level rise over the 1950–2000 period. *J. Clim.* **17**, 2609–2625 (2004).
44. Tsimplis, M. N., Álvarez-Fanjul, E., Gomis, D., Fenoglio-Marc, L. & Pérez, B. Mediterranean Sea level trends: atmospheric pressure and wind contribution. *Geophys. Res. Lett.* **32**, L20602 (2005).
45. Gomis, D., Ruiz, S., Sotillo, M. G., Álvarez-Fanjul, E. & Terradas, J. Low frequency Mediterranean sea level variability: the contribution of atmospheric pressure and wind. *Glob. Planet. Change* **63**, 215–229 (2008).

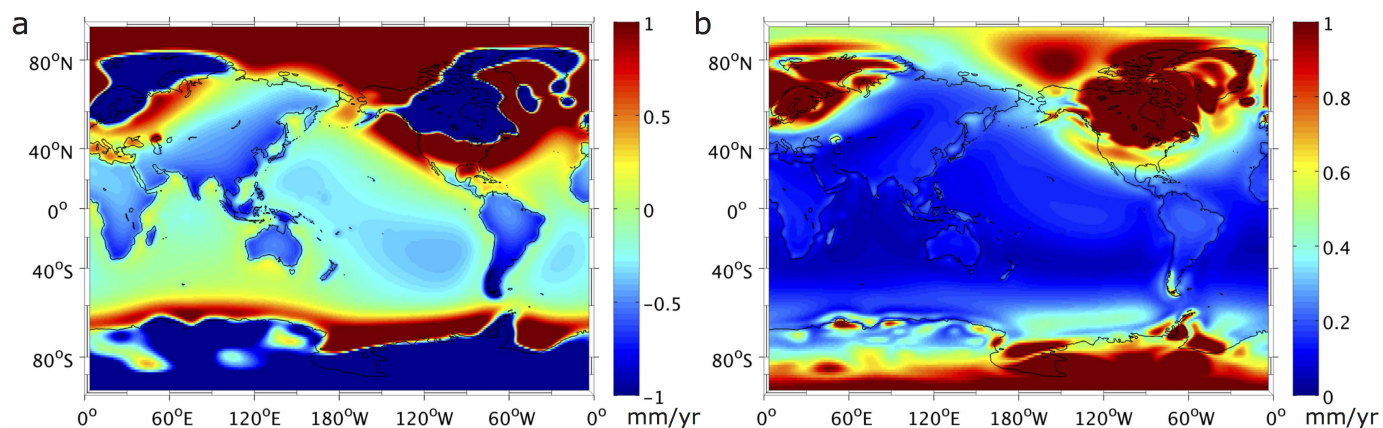
46. Allan, R. & Ansell, T. A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Clim.* **19**, 5816–5842 (2006).
47. Mehra, R. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Automat. Contr.* **15**, 175–184 (1970).
48. Kopp, R. E. *et al.* The impact of Greenland melt on regional sea level: a partially coupled analysis of dynamic and static equilibrium effects in idealized water-hosing experiments. *Clim. Change* **103**, 619–625 (2010).





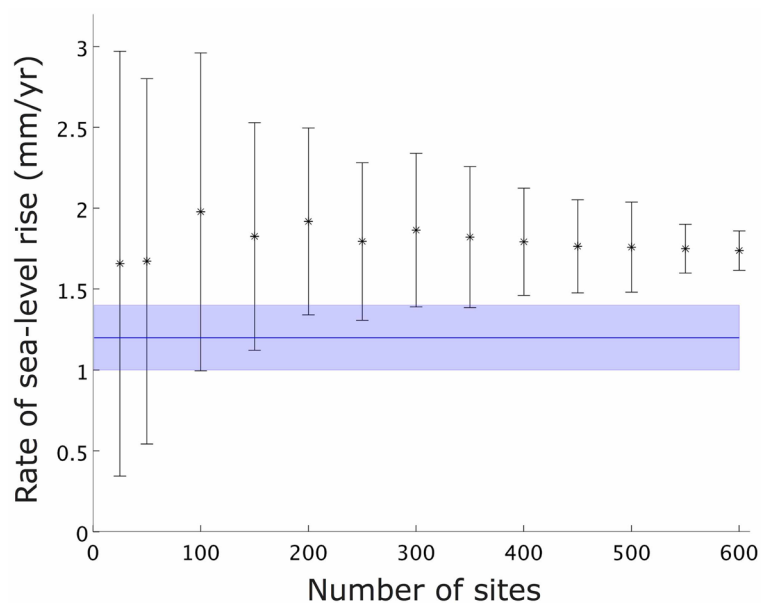
**Extended Data Figure 1 | Illustrative sea-level fingerprints.** **a, b,** Normalized sea-level changes due to rapid melting of the Greenland Ice Sheet (**a**) and the West Antarctic Ice Sheet (**b**). The variable ‘normalized sea-level change’

on the colour scale is formally dimensionless, but may be interpreted as having the unit of metres of sea-level change per metre of the equivalent GMSL change associated with the melt event.



**Extended Data Figure 2 | The present-day rate of change of sea level in  $\text{mm yr}^{-1}$  due to GIA for a suite of Earth models. a, b, Mean sea-level change (a) and standard deviation (b) computed from the output of 161**

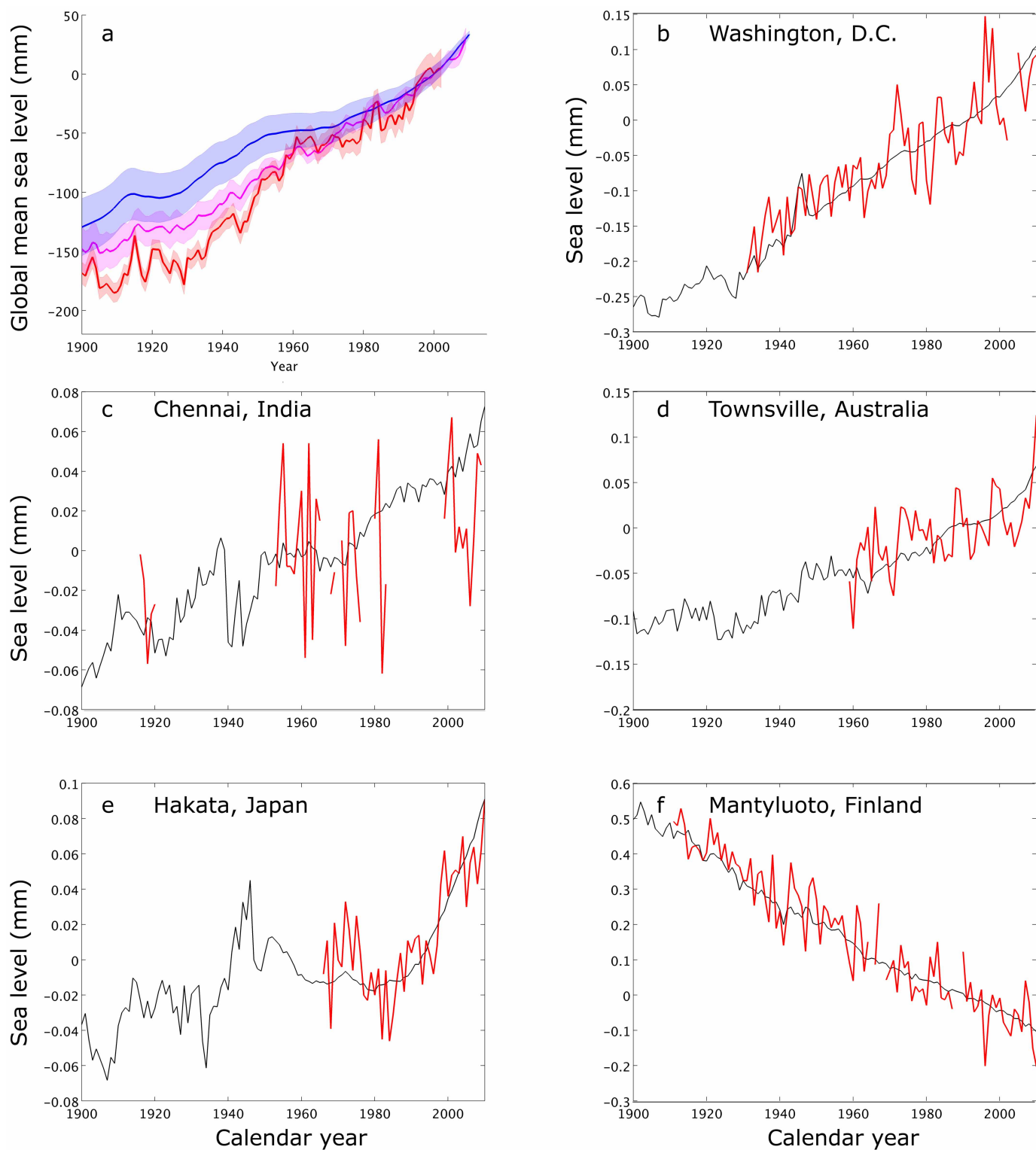
GIA model simulations (see text). In both frames, the colour scale saturates in the near field, which includes areas of post-glacial rebound and peripheral subsidence.



**Extended Data Figure 3 | Bootstrapping analysis of GMSL rate for 1901–90 obtained by sampling the global reconstruction of sea level.** Data points show the mean computed from a bootstrapping analysis of the 1901–90 GMSL rate as a function of the number of geographic sites used in the analysis (ranging from 25 to 600). Error bars,  $\pm 1$  s.d. Sites are obtained by randomly sampling the global KS reconstruction at a subset of tide gauge sites and

introducing data gaps that are consistent with those that exist in the PSMSL database<sup>15</sup>. The analysis was repeated 100 times for each choice of the number of sites. Also shown (horizontal blue line and shading) is the 1901–90 rate and its 90% CI computed from the KS GMSL curve in Fig. 2 ( $1.2 \pm 0.2$  mm yr<sup>-1</sup>; Figs 2 and 3).





**Extended Data Figure 4 | Results of the KS analysis performed using a random subset of 450 tide gauges.** a, KS-estimated GMSL curve derived using a subset of 450 of the 622 tide gauge records discussed in the main text (blue line) and the reconstruction of Church and White<sup>4</sup> (magenta line) and

Jevrejeva *et al.*<sup>3</sup> (red line). The shaded regions represent the 1 $\sigma$  certainty range. Panels b–f show the KS reconstructions (black lines) at a representative set of 5 of the 122 sites that were not used in the estimation procedure. The observations are shown in red.

# A basal ichthyosauriform with a short snout from the Lower Triassic of China

Ryosuke Motani<sup>1</sup>, Da-Yong Jiang<sup>2,3</sup>, Guan-Bao Chen<sup>4</sup>, Andrea Tintori<sup>5</sup>, Olivier Rieppel<sup>6</sup>, Cheng Ji<sup>7</sup> & Jian-Dong Huang<sup>4</sup>

The incompleteness of the fossil record obscures the origin of many of the more derived clades of vertebrates. One such group is the Ichthyopterygia, a clade of obligatory marine reptiles that appeared in the Early Triassic epoch, without any known intermediates<sup>1</sup>. Here we describe a basal ichthyosauriform from the upper Lower Triassic (about 248 million years ago) of China, whose primitive skeleton indicates possible amphibious habits. It is smaller than ichthyopterygians and had unusually large flippers that probably allowed limited terrestrial locomotion. It also retained characteristics of terrestrial diapsid reptiles, including a short snout and body trunk<sup>2</sup>. Unlike more-derived ichthyosauriforms<sup>3</sup>, it was probably a suction feeder. The new species supports the sister-group relationships between ichthyosauriforms and Hupehsuchia<sup>4</sup>, the two forming the Ichthyosauromorpha. Basal ichthyosauromorphs are known exclusively from south China, suggesting that the clade originated in the region, which formed a warm<sup>5,6</sup> and humid<sup>7</sup> tropical archipelago<sup>8</sup> in the Early Triassic. The oldest unequivocal record of a sauropterygian is also from the same stratigraphic unit of the region<sup>9</sup>.

Reptilia Laurenti, 1768

Diapsida Osborn, 1903

Ichthyosauromorpha nov.

**Diagnosis.** The last common ancestor of *Ichthyosaurus communis* and *Hupehsuchus nanchangensis*, and all its descendants. Anterior flanges on humerus and radius present; ulna distal width equal to or greater than proximal width; forelimb longer than or almost equal to hindlimb; manus length at least about three-quarters the length of the stylopodium and zeugopodium combined; fibula extending further post-axially than femur; transverse process of neural arch extremely short or absent.

Ichthyosauriformes nov.

**Diagnosis.** All ichthyosauromorphs more closely related to *Ichthyosaurus communis* than *Hupehsuchus nanchangensis*. Nasal extending anteriorly, well beyond external naris; scleral ring large, filling orbit; snout constricted in dorsal view; converging digits with limited interdigital space.

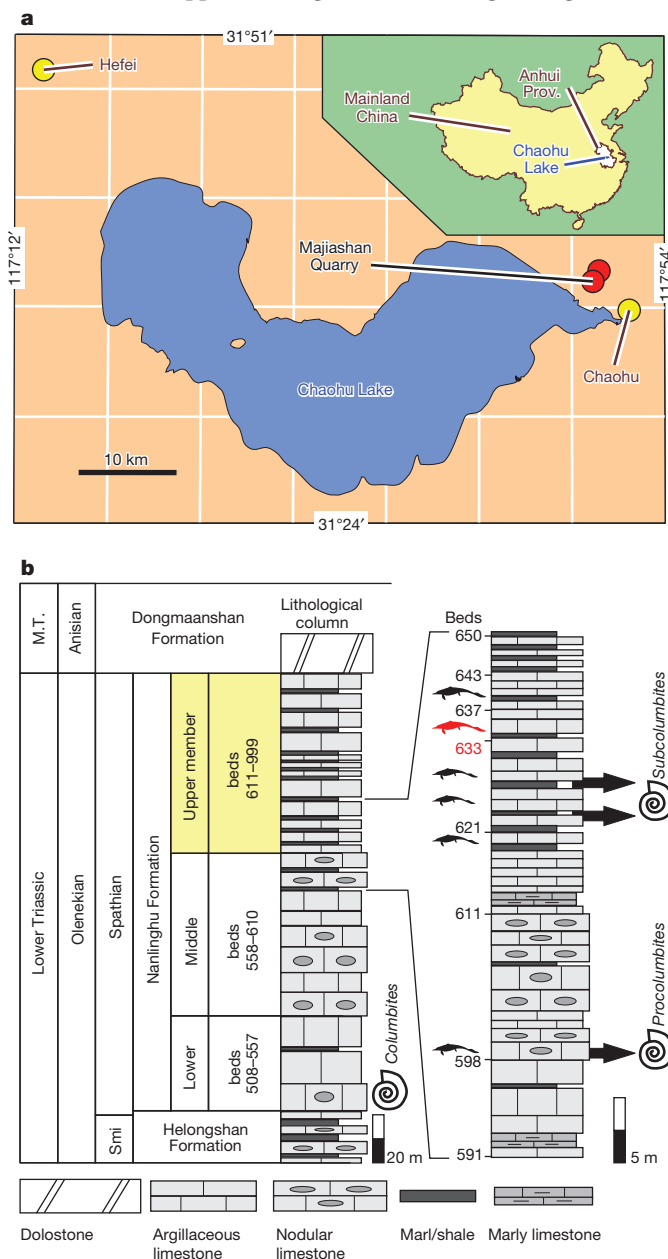
*Cartorhynchus lenticarpus* gen. et sp. nov.

**Etymology.** καρτος (Greek), meaning shortened; ρηψυχος (Greek), meaning snout; lentus (Latin), meaning flexible; carpus (Latin), meaning wrist. Named after truncated snout and cartilaginous wrist.

**Holotype.** Anhui Geological Museum AGB6257.

**Locality and horizon.** From the second level of Majiashan Quarry (31° 37' 26" N, 117° 49' 19" E), near downtown Chaohu, Hefei City, Anhui Province, China (Fig. 1a). Bed 633, about 13 m above the bottom of the Upper Member of the Nanlinghu Formation (Fig. 1b), within the ammonite *Subcolumbites* zone, Spathian, Olenekian, Lower Triassic (Fig. 1b).

**Diagnosis.** Autapomorphies are: snout only half as long as the rest of the skull; very large hyoid; forelimb strongly curved posteriorly; anteriorly curved hindlimb; ribs robust, with proximal intercostal space narrower



**Figure 1 | Locality and horizon of the new species.** **a**, Map of Majiashan relative to Chaohu and Hefei. **b**, Stratigraphic columns of the relevant strata. The specimen is from bed 633 (red) within the Upper Member of the Nanlinghu Formation (yellow). Ichthyopterygian occurrences are indicated by silhouettes. M.T., Middle Triassic; Smi, Smithian. Stratigraphic column in **b** spans from 251.2 to 247.2 million years ago.

<sup>1</sup>Department of Earth and Planetary Sciences, University of California, Davis, One Shields Avenue, Davis, California 95616, USA. <sup>2</sup>Laboratory of Orogenic Belt and Crustal Evolution, Ministry of Education, and Department of Geology and Geological Museum, Peking University, Yiheyuan Street 5, Beijing 100871, China. <sup>3</sup>State Key Laboratory of Palaeobiology and Stratigraphy (Nanjing Institute of Geology and Palaeontology, Chinese Academy of Science), Nanjing 210008, China. <sup>4</sup>Department of Research, Anhui Geological Museum, Jiahe Road 999, Hefei, Anhui 230031, China. <sup>5</sup>Dipartimento di Scienze della Terra, Università degli Studi di Milano, Via Mangiagalli, 34-20133 Milan, Italy. <sup>6</sup>Center of Integrative Research, The Field Museum, Chicago, Illinois 60605-2496, USA. <sup>7</sup>Key Laboratory of Economic Stratigraphy and Palaeogeography, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, 39 East Beijing Road, Nanjing 210008, China.

than ribs; scapula wider distally than proximally; autopodium with broadly spaced tiny ossifications; only three manual digits ossified; gastralia without median element. Other features: mandible deep; trunk shorter than in ichthyopterygians by at least five vertebrae; pineal foramen very large; interclavicle cruciform; parapophyses confluent with anterior vertebral margin.

*Cartorhynchus* is the smallest ichthyosauriform to date. The preserved length of the specimen is 21.4 cm (Fig. 2a). Total body length is estimated to be about 40 cm, assuming tail proportions of basal ichthyopterygians. Of the 31 pre-sacral vertebrae, 5 seem to be cervical. Ichthyopterygians typically have an elongated body with 40–80 pre-sacral vertebrae, except for *Chaohusaurus*, which has about 36 (Fig. 3a). The pre-sacral vertebral count of extant terrestrial reptiles with well-developed limbs ranges from 16 to 36, with 24 being the norm<sup>10</sup>. *Cartorhynchus* is within this terrestrial range (Fig. 3a).

The axial skeleton is heavily built in *Cartorhynchus*. The ribs are pachyostotic, limiting the intercostal space (Fig. 2b). In contrast, the proximal intercostal space is about twice as wide as the ribs in *Chaohusaurus* (Fig. 2c). However, pachyostosis in *Cartorhynchus* is not as pronounced as in the swollen ribs of pachypleurosaurs or sirenians. Osteosclerosis cannot be confirmed without damaging the only specimen. Pachyostosis is a common feature among basal members of secondarily aquatic reptiles<sup>11</sup>. Thickened ribs were lost in Ichthyopterygia but reappeared in the Middle Jurassic *Mollesaurus*<sup>12</sup>.

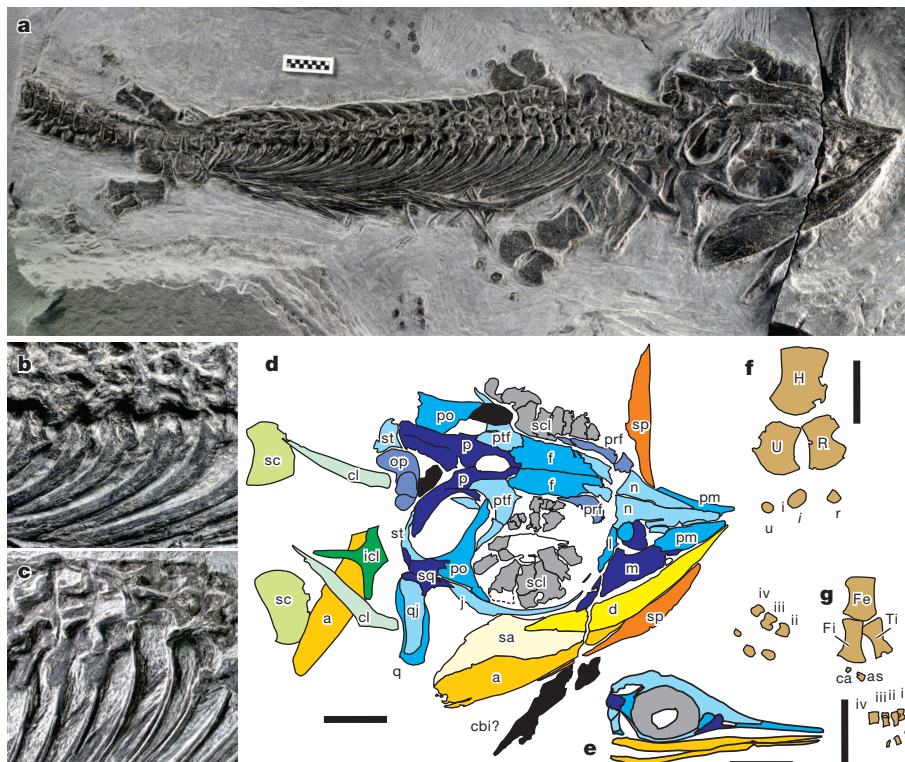
The snout of *Cartorhynchus* is constricted but not elongated (Fig. 2a). Pre-orbital skull length relative to total skull length is shorter in *Cartorhynchus* than in most terrestrial diapsids (Fig. 3b). In contrast, ichthyopterygians uniformly show snout elongation (Fig. 3b). Despite the short snout, the premaxilla of *Cartorhynchus* is elongated, and the external naris is located posteriorly. The nasal of *Cartorhynchus* extends anteriorly to the tip of the snout unlike in most reptiles (Fig. 2d). Elongation

of the nasal and posterior relocation of the external naris preceded snout elongation in ichthyosauriforms.

The short and constricted snout suggests that the gape of this species was much smaller than the mouth cavity, enabling pressure concentration for suction feeding<sup>3</sup>. The inference of suction feeding in *Cartorhynchus* is further supported by the presence of a large and robust hyobranchial element (Fig. 2a, d), and also by edentulism. There is, however, a dental groove at least on the right mandibular ramus, and this is also present in the edentulous *Hupehsuchus*<sup>13</sup>.

The distance between the zeugopodium and metapodium is unusually large, suggesting the presence of cartilaginous centralia. These elements only occur in *Chaohusaurus* among ichthyopterygians<sup>14</sup>. The carpal gap in *Cartorhynchus* is even wider than in a specimen of *Chaohusaurus* in which both lateral and medial centralia are ossified<sup>14</sup> (Fig. 3c). We re-prepared both limbs twice but did not find any additional elements.

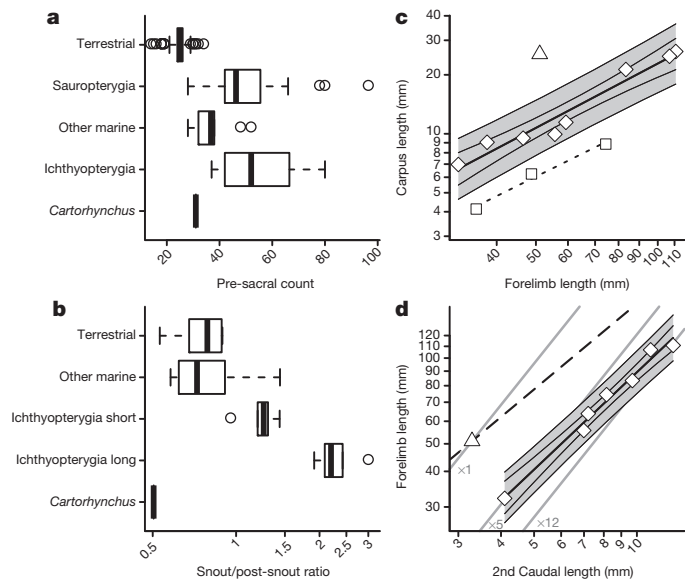
The flippers may well have allowed limited terrestrial locomotion given their unusually large size (Fig. 3d). The carpus allowed dorso-ventral flexion of the flipper without a functional elbow, enabling seal-like flipper bending on land—flipper bendability is essential in terrestrial locomotion<sup>15</sup>. The cartilaginous forelimb may not have been as strong as fully ossified flippers but the mass it supported was small. The small body size, through scaling effects, provided a low body mass/flipper area ratio that was about one-fifth to one-twelfth of the values in *Chaohusaurus*, assuming similar body shapes (Fig. 3d, grey lines). Note that juvenile sea turtles, with highly cartilaginous wrist and ankle joints within the flipper, are able to undertake safe locomotion on land. Considering that the pachyostotic ribs would ballast the body in surging water near shores, *Cartorhynchus* may have been amphibious. Added mass from thick ribs may be disadvantageous for terrestrial locomotion but they are not as heavy as turtle shells. Moreover, even dugongs, which are heavily pachyostotic, are known to occasionally give birth on the beach<sup>16</sup>. A short trunk and



**Figure 2 | The holotype of *Cartorhynchus lenticarpus* gen. et sp. nov.**  
**a**, Whole specimen. **b**, Close-up of ribs. **c**, Ribs of *Chaohusaurus* (AGM CH-628-16) for comparison. **d**, Skull and shoulder elements. **e**, Skull of a newborn *Chaohusaurus*<sup>20</sup> drawn to the same scale as **d** for comparison. **f**, Right forelimb. **g**, Right hindlimb. a, angular; as, astragalus; ca, calcaneum; cbi, first ceratobranchial; cl, clavicle; d, dentary; f, frontal; Fe, femur; Fi, fibula;

H, humerus; i, intermedium; icl, interclavicle; j, jugal; l, lacrimal; m, maxilla; n, nasal; op, opisthotic; p, parietal; pm, prefrontal; po, postorbital; prf, prefrontal; ptf, postfrontal; q, quadrate; qj, quadratojugal; R, radius; r, radiale; sa, surangular; sc, scapula; scl, scleral ossicles; sp, splenial; sq, squamosal; st, supratemporal; Ti, tibia; U, ulna; u, ulnare; i–v (in **g**), metapodials. Scale bars, 1 cm.





**Figure 3 | Quantitative comparisons of selected characteristics.** **a**, Pre-sacral count. **b**, Snout/post-snout ratio of the skull. **c**, Carpus/whole-limb ratio of the forelimb. **d**, Forelimb length versus the second caudal centrum length, a proxy for body size. In **c** and **d**, triangles are *Cartorhynchus*, squares are *Chaohusaurus geishanensis* and diamonds are *C. chaioxianensis*. Dashed line in **d** represents the same forelimb/vertebra proportion as in *Cartorhynchus*. Grey lines in **d** are isoclines of the body mass/flippers area ratio, assuming that the cube of the second caudal centrum length is approximately proportional to body mass, and the length of the flipper is roughly proportional to its area. Symbols of box plots in **a** and **b**: circle, outlier; rectangle, 50 percentile; thick line, median; whisker, extreme point or, when there are outliers, data range times interquartile range. See refs 13 and 14 for data sources.

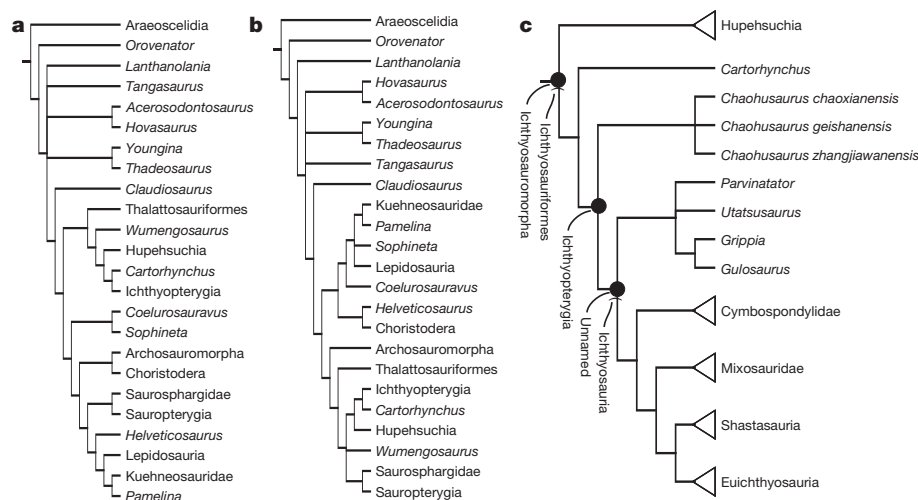
snout would have been advantageous in terrestrial locomotion. The forelimb flipper is strongly curved, with the phalangeal axis tilted about 50° post-axially relative to the zeugopodial axis. Both right and left limbs display similar angles, suggesting that the curvature is natural. The curvature allowed the tip of the flipper to be kept close to the body, raising the mechanical advantage of terrestrial body support by flippers.

With a well-ossified axial skeleton and finished surfaces on many skull bones, the holotype appears to be mature, but its appendicular skeleton is poorly ossified. Given the large size of the forelimb, it seems likely that

its ossification was delayed through paedomorphosis, as is the case in many basal marine reptiles<sup>4,14,17–19</sup>. We consider the holotype to be almost, if not fully, mature. However, with only one specimen available, the possibility of immaturity cannot be completely rejected. It is unlikely that *Cartorhynchus* is a young individual of a known species. Phylogenetic characters eliminate non-ichthyosauriform diapsids from such a comparison (Fig. 4), while a list of features prevents it from being an immature ichthyopterygian. The snout of ichthyopterygians is already elongated in newborns<sup>20,21</sup> (for example, Fig. 2e), unlike in *Cartorhynchus* (Fig. 2d). The pre-sacral vertebral count is not expected to increase after birth<sup>10</sup>. Furthermore, only adult ichthyopterygians have forelimbs that are almost as long as the skull, as in *Cartorhynchus*. Additional features may be compared to *Chaohusaurus*, the only ichthyopterygian genus that co-occurs with the new species. The mandible of *Cartorhynchus* (55.4 × 12.2 mm, length by depth) is distinctively deeper than that of AGM-CH-628-19 (103.4 × 7.0 mm), the largest individual of *Chaohusaurus chaioxianensis*. The ceratobranchial is very large (22.8 × 4.0 mm) unlike in AGM-CH-628-19 (11.7 × 1.28 mm), and the radiale is already ossified, a bone that appears only in the largest individual of *C. chaioxianensis*<sup>14,19</sup>.

The eye appears large relative to the skull largely because the snout is short; however, it is not proportionally larger than in some terrestrial reptiles, such as araucoscelidians. Three characteristics of the scleral ring are indicative of visual optics, namely the absolute size that determines the number of retinal cells, absolute aperture size that determines the ability to spot bioluminescent prey in the dark, and relative aperture size that is related to *f* numbers<sup>22</sup>, an optical index that describes relative brightness. For all three features *Cartorhynchus* has the smallest values of ichthyosauriforms, indicating a lesser degree of adaptation to underwater vision.

The phylogeny shows Hupehsuchia<sup>13</sup>, *Cartorhynchus* and *Chaohusaurus*, all known exclusively from the south China block, to be the three most basal members of Ichthyosauromorpha (Fig. 4). The oldest ichthyosauromorph fossil is from the *Procolymbites* zone of the Spathian (Lower Triassic) of Chaohu<sup>20</sup> (Fig. 1b), but many more are known from the overlying *Subcolymbites* zone at multiple locations around the Panthalassa<sup>1</sup>. Marine reptiles older than the *Procolymbites* zone are unknown despite pertinent outcrops around the world. Multiple exposures of relevant strata are accessible in Chaohu, yet persistent prospecting efforts have failed to locate a reptile fossil in underlying Spathian rocks, or in Smithian strata rich in fossil fishes (Fig. 1b). Furthermore, decades of limestone mining activities in the region did not uncover any older reptile fossils. Their absence coincides with the recent report that *Chaohusaurus* from the *Subcolymbites* zone still retained a terrestrial



**Figure 4 | Phylogenetic hypotheses of *Cartorhynchus*.** **a**, Position of *Cartorhynchus* among Diapsida, with aquatic adaptations excluded. **b**, As in **a** but with aquatic adaptations included. **c**, Relationships of basal ichthyosauromorphs. Analyses are based on most recent phylogenetic data sets

for marine reptile relationships<sup>4</sup> and ichthyopterygian phylogeny (C. Ji *et al.*, manuscript in preparation), respectively (Supplementary Information). All trees were abbreviated from more complete topologies (Extended Data Figs 1–3).

mode of reproduction that carried a risk in water<sup>20</sup>, suggesting that ichthyosauromorphs invaded the sea not much earlier than the appearance of *Chaohusaurus*. A recent statement that ichthyosauromorphs were present in the Smithian<sup>5</sup> is a misunderstanding based on outdated references<sup>23,24</sup> that have been updated<sup>25,26</sup>.

The causes driving marine invasion could be multiple, including predation pressure and competition for food that may be lower in the sea than on adjacent land<sup>27,28</sup>. Notably, Ichthyosauromorpha was not the only marine reptile group to have emerged in the Early Triassic: the oldest unequivocal record of Sauropterygia is also from Chaohu<sup>9</sup>, 45 beds above the oldest ichthyopterygian fossil. The south China block was in the tropical latitudes at the time<sup>29</sup>, forming a warm<sup>5,6</sup> and humid<sup>7</sup> archipelago<sup>8</sup>. Future studies would be required to test if any climatic and geographic factors may have encouraged marine invasion.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 10 June; accepted 15 September 2014.**

**Published online 5 November 2014.**

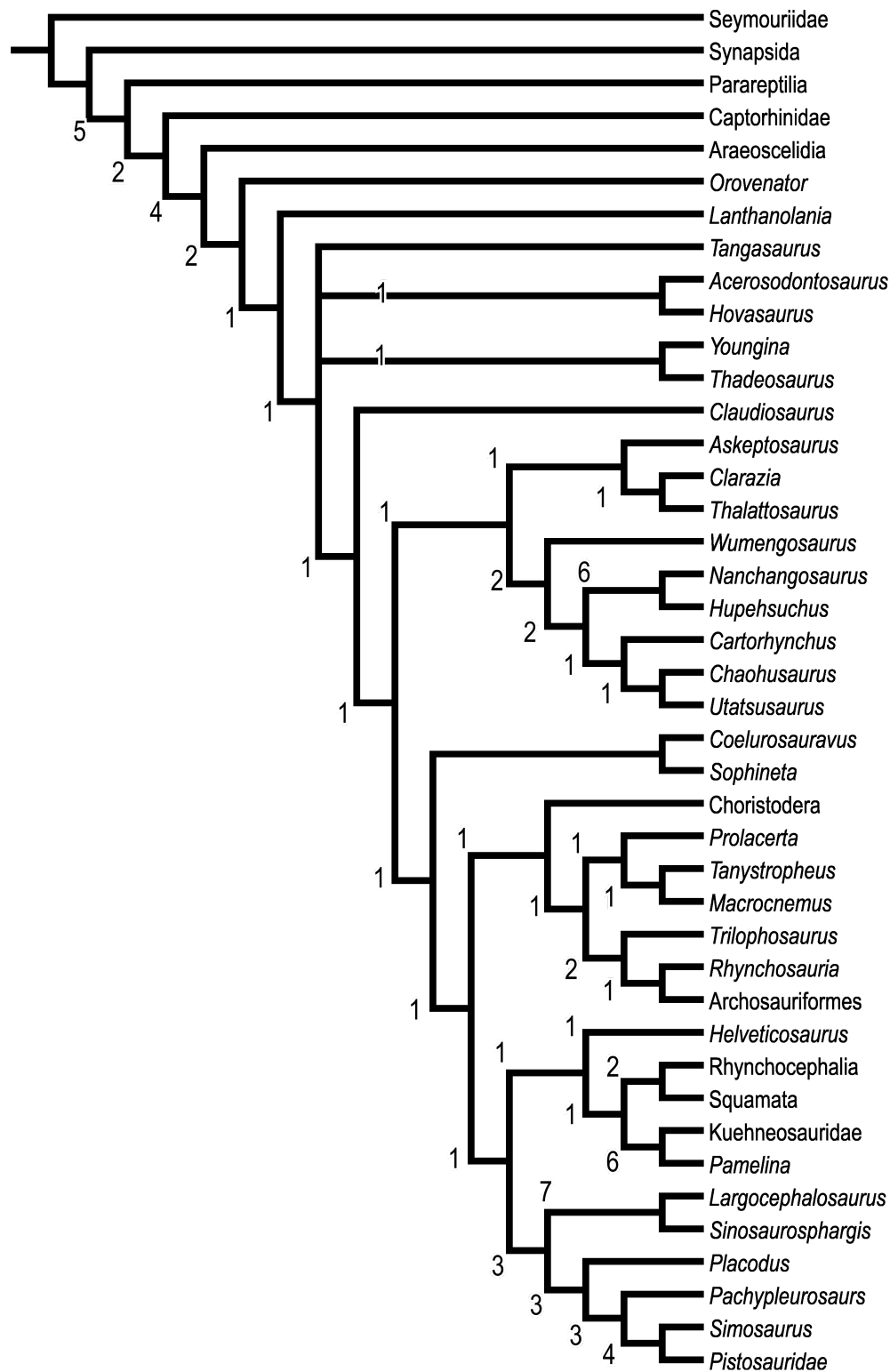
- McGowan, C. & Motani, R. *Ichthyopterygia* Vol. 8 (Verlag Dr. Friedrich Pfeil, 2003).
- Müller, J. *et al.* Homeotic effects, somitogenesis and the evolution of vertebral numbers in recent and fossil amniotes. *Proc. Natl Acad. Sci. USA* **107**, 2118–2123 (2010).
- Motani, R. *et al.* Absence of suction feeding ichthyosaurs and its implications for Triassic Mesopelagic paleoecology. *PLoS ONE* **8**, e66075 (2013).
- Chen, X.-h., Motani, R., Cheng, L., Jiang, D.-y. & Rieppel, O. The enigmatic marine reptile *Nanchangosaurus* from the Lower Triassic of Hubei, China and the phylogenetic affinity of *Hupehsuchus*. *PLoS ONE* **9**, e102361 (2014).
- Sun, Y. D. *et al.* Lethally hot temperatures during the Early Triassic greenhouse. *Science* **338**, 366–370 (2012).
- Goudemand, N., Romano, C., Brayard, A., Hochuli, P. A. & Bucher, H. Comment on “Lethally Hot Temperatures During the Early Triassic Greenhouse”. *Science* **339**, 1033 (2013).
- Péron, S., Bourquin, S., Fluteau, F. & Guillocheau, F. Paleoenvironment reconstructions and climate simulations of the Early Triassic: impact of the water and sediment supply on the preservation of fluvial systems. *Geodin. Acta* **18**, 431–446 (2005).
- Xiao, W. J. & He, H. Q. Early Mesozoic thrust tectonics of the northwest Zhejiang region (Southeast China). *Geol. Soc. Am. Bull.* **117**, 945–961 (2005).
- Jiang, D. *et al.* Early Triassic eosauropterygian *Majiasuchosaurus discocoracoides*, gen. et sp. nov. (Reptilia, Sauropterygia) from Chaohu, Anhui Province, China. *J. Vertebr. Paleontol.* **34**, 1044–1052 (2014).
- Hoffstetter, R. & Gasc, J.-P. in *Biology of the Reptilia*. Vol 1 (eds Gans, C., Bellairs, A. d'A. & Parsons, T. S.) 201–310 (Academic, 1969).
- Houssaye, A. “Pachyostosis” in aquatic amniotes: a review. *Integr. Zool.* **4**, 325–340 (2009).
- Talevi, M. & Fernandez, M. S. Unexpected skeletal histology of an ichthyosaur from the Middle Jurassic of Patagonia: implications for evolution of bone microstructure among secondary aquatic tetrapods. *Naturwissenschaften* **99**, 241–244 (2012).
- Carroll, R. L. & Dong, Z. *Hupehsuchus*, an enigmatic aquatic reptile from the Triassic of China, and the problem of establishing relationships. *Phil. Trans. R. Soc. B* **331**, 131–153 (1991).
- Motani, R. *et al.* First evidence of centralia in Ichthyopterygia reiterating bias from pedomorphic characters on marine reptile phylogenetic reconstruction. *J. Vertebr. Paleontol.* (in the press).
- Mazouchova, N., Umbanhowar, P. B. & Goldman, D. I. Flipper-driven terrestrial locomotion of a sea turtle-inspired robot. *Bioinspir. Biomim.* **8**, 026007 (2013).
- Marsh, H., Heinsohn, G. E. & Marsh, L. M. Breeding cycle life history and population dynamics of the *Dugong Dugong-Dugon Sirenia Dugongidae*. *Aust. J. Zool.* **32**, 767–788 (1984).
- Rieppel, O. *Helveticosaurus zollingeri* Peyer (Reptilia Diapsida) skeletal pedomorphosis, functional anatomy and systematic affinities. *Palaeontographica Abteilung A* **208**, 123–152 (1989).
- Caldwell, M. W. Limb ontogeny, evolution and aquatic adaptation in lepidosauromorph diapsids. *J. Vertebr. Paleontol.* **14**, 19A (1994).
- Motani, R. *et al.* Status of *Chaohusaurus chaoxianensis* (Young and Dong, 1972). *J. Vertebr. Paleontol.* (in the press).
- Motani, R., Jiang, D., Tintori, A., Rieppel, O. & Chen, G. B. Terrestrial origin of viviparity indicated by the oldest embryonic fossil of Mesozoic marine reptiles. *PLoS One* **9**, e8B640 (2014).
- Boettcher, R. New information on the reproductive biology of Ichthyosaurs (Reptilia). *Stuttgarter Beitrage zur Naturkunde Serie B* **164**, 1–51 (1990).
- Motani, R., Rothschild, B. M. & Wahl, W. Large eyeballs in diving ichthyosaurs. *Nature* **402**, 747 (1999).
- Callaway, J. M. & Brinkman, D. B. Ichthyosaurs (Reptilia, Ichthyosauria) from the Lower and Middle Triassic Sulfur Mountain Formation, Wapiti Lake Area, British-Columbia, Canada. *Can. J. Earth Sci.* **26**, 1491–1500 (1989).
- Cox, C. B. & Smith, D. G. Review of Triassic vertebrate faunas of Svalbard. *Geol. Mag.* **110**, 405–418 (1973).
- Nicholls, E. L. & Brinkman, D. B. in *Vertebrate Fossils and the Evolution of Scientific Concepts* (ed. Sarjeant, W. A. S.) 521–535 (Gordon and Breach, 1995).
- Harland, W. B. The geology of svalbard. *Mem. Geol. Soc. Lond.* **17**, 1–521 (1997).
- Carroll, R. L. Evolutionary constraints in aquatic diapsid reptiles. *Spec. Pap. Palaeontol.* **33**, 145–155 (1985).
- Vermeij, G. J. & Dudley, R. Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biol. J. Linn. Soc.* **70**, 541–554 (2000).
- Sun, Z. M. *et al.* Magnetostratigraphy of the Lower Triassic beds from Chaohu (China) and its implications for the Induan-Olenekian stage boundary. *Earth Planet. Sci. Lett.* **279**, 350–361 (2009).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank T. Sato for the excellent preparation of the holotype. The study was enabled by grants from the National Geographic Society Committee for Research and Exploration (8669-09) to R.M., Project 40920124002 and 41372016 from the National Natural Science Foundation of China to D.-Y.J., Project 123102 from State Key Laboratory of Palaeobiology and Stratigraphy (Nanjing Institute of Geology and Palaeontology, CAS) to D.-Y.J., Project 20120001110072 from the Research Fund for the Doctoral Program of Higher Education to D.-Y.J., and a Project of Protection for Geological Heritage from Department of Land and Resource of Anhui Province to G.-B.C.

**Author Contributions** R.M. conceived the study, participated in the relevant fossil excavations, ran all analyses, drew all figures except Fig. 1b, and wrote the manuscript. D.-Y.J. conceived the study, supervised the relevant fossil excavations and preparations, drew Fig. 1b, and revised the manuscript; G.-B.C. conceived the study and supervised off-season fossil collections; A.T. conceived the study, participated in the relevant fossil excavations, and revised the manuscript; O.R. conceived the study, participated in the relevant fossil excavations, and revised the manuscript; C.J. provided an unpublished data matrix, participated in the relevant fossil excavations, and revised the manuscript; J.-D.H. helped supervise off-season fossil collections.

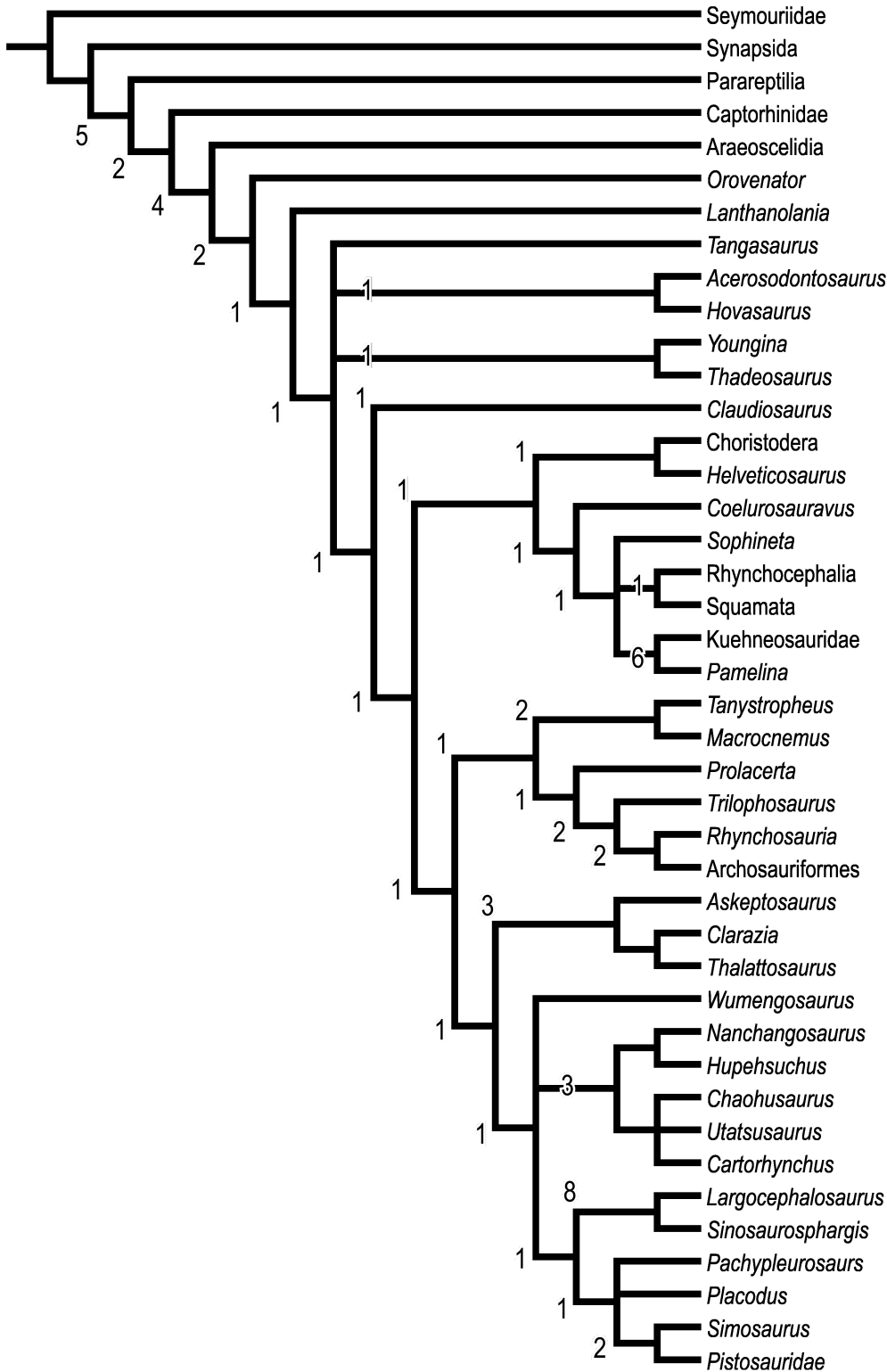
**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.M. (rmotani@ucdavis.edu) or D.-Y.J. (djiang@pku.edu.cn).



**Extended Data Figure 1 | Phylogenetic hypothesis of diapsid relationships when aquatic adaptations are recoded as ambiguous.** Strict consensus of two most parsimonious trees (tree length (TL) = 805, consistency index (CI) = 0.317, retention index (RI) = 0.587) obtained by a heuristic search in

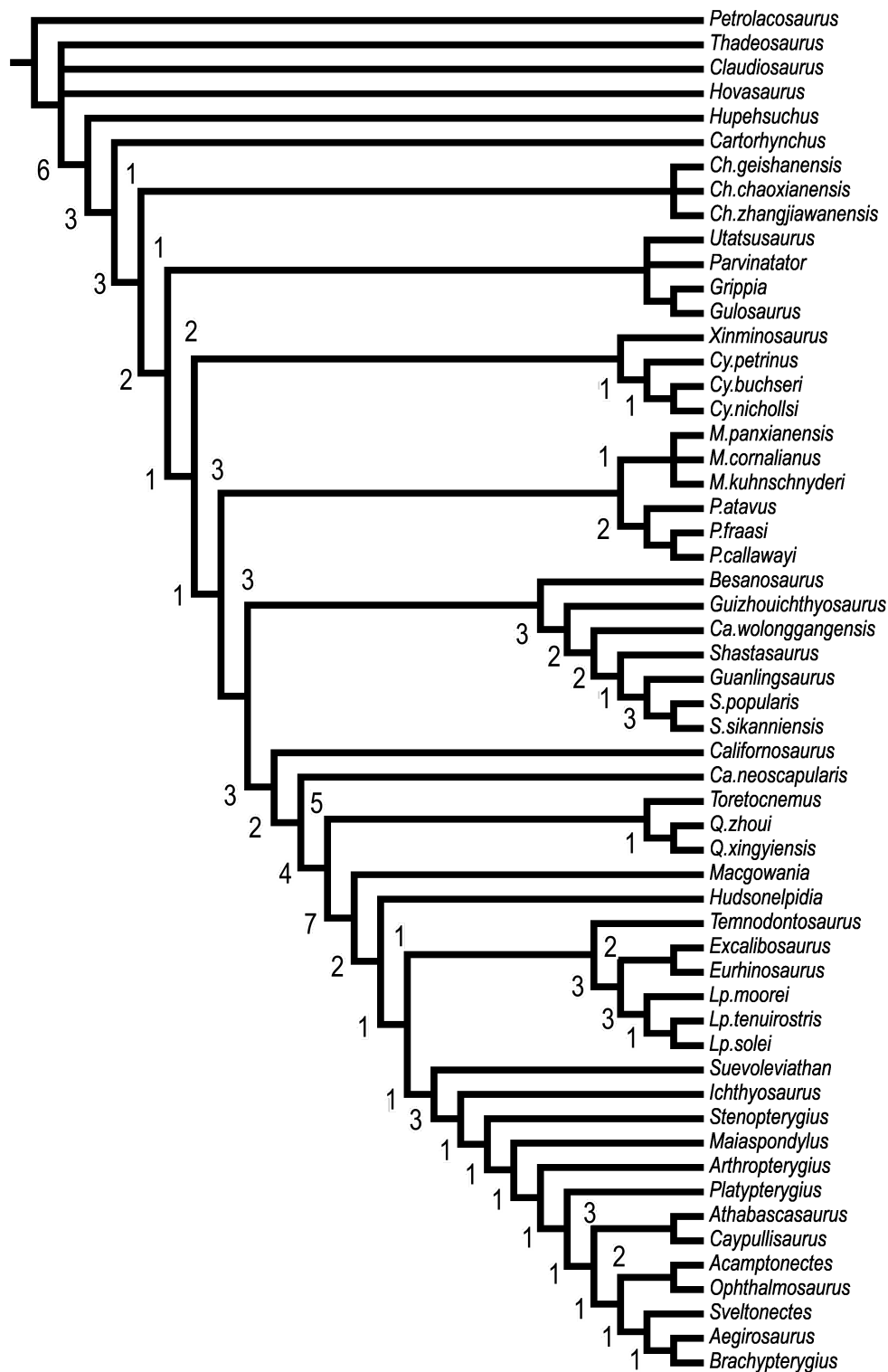
PAUP\* 4b10 (hold = 10, nreps = 100, addseq = random, swap = tbr). Numbers associated with clades are Bremer support values calculated in TNT 1.1. Original tree of Fig. 4a. Phylogenetic tree is based on the data described in Supplementary Information.





**Extended Data Figure 2 | Phylogenetic hypothesis of diapsid relationships when aquatic adaptations are recoded normally.** Strict consensus of 20 most parsimonious trees (TL = 839, CI = 0.311, RI = 0.612) obtained by a heuristic search in PAUP\* 4b10 (hold = 10, nreps = 100, addseq = random,

swap = tbr). Numbers associated with clades are Bremer support values calculated in TNT 1.1. Original tree of Fig. 4b. Phylogenetic tree is based on the data described in Supplementary Information.



**Extended Data Figure 3 | Phylogenetic hypothesis of Ichthyosauriformes.** Strict consensus of 82 most parsimonious trees (TL = 527, CI = 0.423, RI = 0.796) obtained by a heuristic search in PAUP\* 4b10 (hold = 10,

nreps = 100, addseq = random, swap = tbr). Numbers associated with clades are Bremer support values calculated in TNT 1.1. Original tree of Fig. 4c. Phylogenetic tree is based on the data described in Supplementary Information.

# The mutational landscapes of genetic and chemical models of *Kras*-driven lung cancer

Peter M. K. Westcott<sup>1,2</sup>, Kyle D. Halliwill<sup>1,2</sup>, Minh D. To<sup>1</sup>, Mamunur Rashid<sup>3</sup>, Alistair G. Rust<sup>3</sup>, Thomas M. Keane<sup>3</sup>, Reyno Delrosario<sup>1</sup>, Kuang-Yu Jen<sup>4</sup>, Kay E. Gurley<sup>5</sup>, Christopher J. Kemp<sup>5</sup>, Erik Fredlund<sup>6</sup>, David A. Quigley<sup>1</sup>, David J. Adams<sup>3</sup> & Allan Balmain<sup>1,7</sup>

Next-generation sequencing of human tumours has refined our understanding of the mutational processes operative in cancer initiation and progression, yet major questions remain regarding the factors that induce driver mutations and the processes that shape mutation selection during tumorigenesis. Here we performed whole-exome sequencing on adenomas from three mouse models of non-small-cell lung cancer, which were induced either by exposure to carcinogens (methyl-nitrosourea (MNU) and urethane) or by genetic activation of *Kras* (*Kras*<sup>LA2</sup>). Although the MNU-induced tumours carried exactly the same initiating mutation in *Kras* as seen in the *Kras*<sup>LA2</sup> model (G12D), MNU tumours had an average of 192 non-synonymous, somatic single-nucleotide variants, compared with only six in tumours from the *Kras*<sup>LA2</sup> model. By contrast, the *Kras*<sup>LA2</sup> tumours exhibited a significantly higher level of aneuploidy and copy number alterations compared with the carcinogen-induced tumours, suggesting that carcinogen-induced and genetically engineered models lead to tumour development through different routes. The wild-type allele of *Kras* has been shown to act as a tumour suppressor in mouse models of non-small-cell lung cancer. We demonstrate that urethane-induced tumours from wild-type mice carry mostly (94%) *Kras* Q61R mutations, whereas those from *Kras* heterozygous animals carry mostly (92%) *Kras* Q61L mutations, indicating a major role for germline *Kras* status in mutation selection during initiation. The exome-wide mutation spectra in carcinogen-induced tumours overwhelmingly display signatures of the initiating carcinogen, while adenocarcinomas acquire additional C > T mutations at CpG sites. These data provide a basis for understanding results from human tumour genome sequencing, which has identified two broad categories of tumours based on the relative frequency of single-nucleotide variations and copy number alterations<sup>1</sup>, and underline the importance of carcinogen models for understanding the complex mutation spectra seen in human cancers.

Sequencing studies of human cancers have identified a number of mutation 'signatures', suggesting that tumours carry an imprint of the environmental agents to which patients were exposed<sup>2–4</sup>. At present, there are no studies of genome-wide carcinogen signatures in any mouse cancer models, despite widespread use of these models in studies of cancer. To address the importance of engineered versus carcinogen-induced mutations, we investigated the mutations in mouse non-small-cell lung carcinoma (NSCLC) arising as a result of spontaneous oncogenic activation of *Kras* (*Kras*<sup>LA2</sup>)<sup>5</sup>, or exposure to urethane or MNU<sup>6</sup>. Both carcinogens initiate lung tumorigenesis by oncogenic mutation of *Kras*, which is frequently mutated in human NSCLC<sup>7</sup>. Whole-exome sequencing (WES) was performed on 82 lung adenomas from FVB/N mice, 44 induced by urethane, 26 by MNU, and 12 by the *Kras*<sup>LA2</sup> allele (Extended Data Table 1). To study the tumour suppressive role of wild-type *Kras*, we included mice with one functionally null *Kras* allele, *Kras*<sup>+/-LSL-G12D</sup>

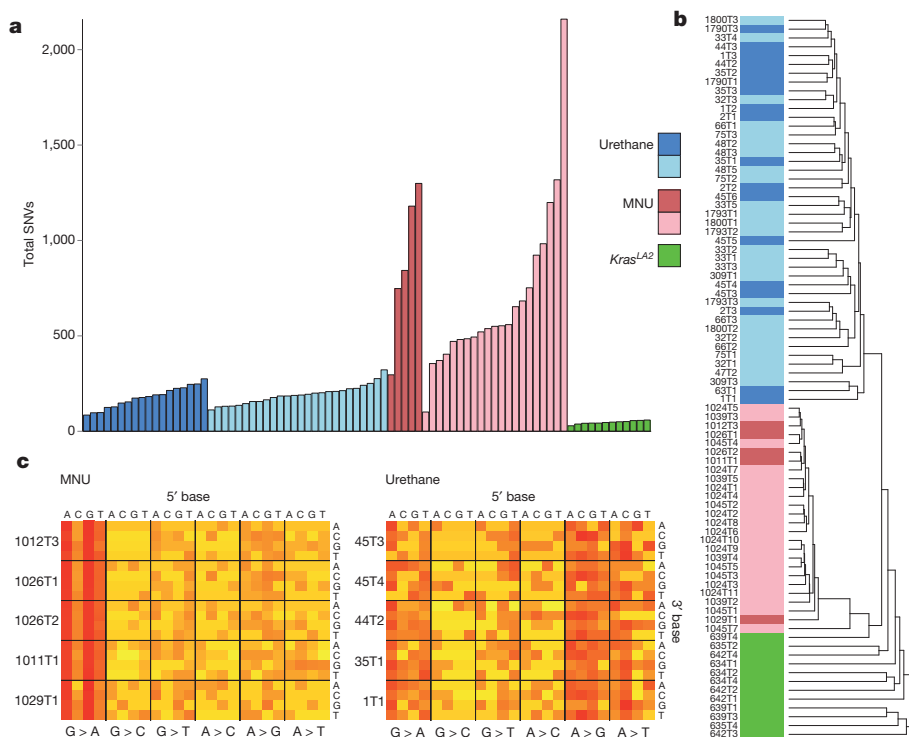
(see Methods)<sup>8</sup>, hereafter referred to as *Kras*<sup>+/-</sup>. Importantly, these mice develop more and larger tumours than wild-type littermates after carcinogen treatment<sup>9,10</sup>.

Carcinogen-induced tumours had far more single-nucleotide variations (SNVs) than *Kras*<sup>LA2</sup> tumours (Fig. 1a), with an average of 728 and 185 in MNU- and urethane-induced tumours, respectively, and 47 in *Kras*<sup>LA2</sup> tumours. This is similar to findings in humans, in which lung tumours from smokers contained orders of magnitude more SNVs than tumours from non-smokers<sup>11</sup>. We performed hierarchical clustering on the 96 possible SNVs, classified by trinucleotide context and substitution<sup>3</sup>, and found that tumours cluster perfectly by treatment (Fig. 1b), underscoring distinct mutational spectra. Highly consistent signatures are apparent across all tumours of each carcinogen group (Fig. 1c and Extended Data Fig. 1a, b), in agreement with the known A > T, A > G, and G > A substitutions induced by urethane<sup>12</sup>, and G > A transitions induced by MNU<sup>13</sup>. The elevated SNV burden and clear carcinogen imprint show that most SNVs were induced during the period of carcinogen activity after administration. By contrast, *Kras*<sup>LA2</sup> tumours showed no notable signatures (Extended Data Fig. 1c).

A highly significant 5'-flanking purine bias and 3'-flanking thymidine bias for G > A transitions was identified in the MNU-induced tumours (Extended Data Fig. 2a). Indeed, GGT > A is the most frequent SNV in this group. In urethane-induced tumours, a slight bias for 3'-cytidine in A > G transitions and 3'-guanine in A > T transversions was seen (Extended Data Fig. 2b, c), while G > A transitions were also common (Extended Data Fig. 2d). The most frequent SNVs in *Kras*<sup>LA2</sup> tumours were CGN > A (or the complement, NCG > T) (Extended Data Fig. 2e). Importantly, CGN > A is a signature of genomic instability in cancer and normal cells<sup>3,14,15</sup>.

In concordance with the propensity of MNU to induce GGT > A transitions, 25/26 MNU-induced lung tumours carried this transition in codon 12 of *Kras*, resulting in a G12D mutation, while all 44 urethane-induced tumours harboured *Kras* Q61 mutations (Supplementary Table 1). Histological evaluation revealed the expected tumour types (Extended Data Fig. 3a), and solid tumours were significantly enriched in the MNU and *Kras*<sup>LA2</sup> groups, which share the *Kras* G12D mutation (Extended Data Fig. 3b). It is possible that *Kras* G12D initiates a pathway to solid NSCLC that is distinct from that initiated by Q61 mutants. Alternatively, urethane may induce *Kras* mutations in a different population of tumour-initiating cells. Remarkably, urethane-induced tumours from wild-type mice had almost exclusively *Kras* Q61R mutations, whereas tumours from *Kras*<sup>+/-</sup> mice had almost exclusively Q61L mutations (Extended Data Fig. 4a, b). This switch is probably not due to differences in carcinogen metabolism or DNA repair, as neither the overall mutation spectra (Extended Data Fig. 1b) nor the exome-wide rates of the causative Q61R and Q61L substitutions (Extended Data Fig. 4c) differed between tumours of the two genotypes. This suggests that *Kras*

<sup>1</sup>Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California 94158, USA. <sup>2</sup>Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, California 94158, USA. <sup>3</sup>Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1HH, UK. <sup>4</sup>Department of Pathology, University of California San Francisco, San Francisco, California 94143, USA. <sup>5</sup>Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. <sup>6</sup>Science for Life Laboratory, Department of Oncology-Pathology, Karolinska Institute, Stockholm 171 21, Sweden. <sup>7</sup>Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94158, USA.



**Figure 1 | Differences in mutation burden and spectra between carcinogen and genetic models.** **a**, Total SNVs per tumour. Light shades denote *Kras*<sup>+/−</sup> genotype. All comparisons of SNVs between treatment groups were significant ( $P \leq 1.0 \times 10^{-6}$ , Wilcoxon rank-sum test, Holm's correction). No significant differences were observed between wild-type and *Kras*<sup>+/−</sup> tumours. **b**, Unsupervised, hierarchical clustering of tumours by trinucleotide context substitutions. **c**, Stacked heatmaps of mutation spectra for five representative MNU-induced and urethane-induced tumours (see Extended Data Fig. 1 for all tumours). Substitutions are shown below each heatmap, with the 5'- and 3'-flanking base displayed on top and right, respectively.

Q61R and Q61L are functionally distinct, and that the selection of cells harbouring these oncoproteins is modulated by wild-type *Kras*. Intriguingly, in the single instance of a *Kras* Q61L mutant tumour from a wild-type mouse, a *Kras* loss-of-function mutation (T35A)<sup>16,17</sup> was also found, potentially inactivating the wild-type allele. Although *KRAS* Q61 mutations are relatively rare in human lung cancer, further investigation of the Q61 switch may yield valuable insights into *RAS* mutation selection, and the interplay of *RAS* oncogenes and their proto-oncogenes. Although further studies are needed to identify the mechanism of this selection, we conclude that *Kras* mutations are not only carcinogen-dependent, but are influenced by germline differences that alter the expression of wild-type *Kras*.

We focused our search for additional driver mutations on genes known to harbour bona fide driver mutations in human cancers<sup>18,19</sup> (see Methods). Sixty-five consequential SNVs in 49 of these genes were validated (Extended Data Table 2), most involving amino acids conserved between mouse and human. SNVs in *Akt1*, *Atm*, *Rnf43*, *Notch1*, *Ret* and *Rb1*, in particular, occurred at positions homologous to mutations in human cancers (Supplementary Table 2). Two nonsense and two missense mutations were found in *Mtus1*, a candidate tumour suppressor gene in multiple cancers<sup>20–23</sup>. In concordance with its role as a tumour suppressor, knockdown of *Mtus1* accelerated growth in a mouse lung cancer cell line driven by *Kras* G12D (Extended Data Fig. 5a, b). In addition, *MTUS1* expression is significantly and positively associated with overall survival across all stages in human lung adenocarcinoma (The Cancer Genome Atlas (TCGA) Human Lung Adenocarcinoma (LUAD) data set RNA-seq,  $n = 354$ ) (Extended Data Fig. 5c and Supplementary Table 3). This association was validated in an independent human lung adenocarcinoma data set<sup>24</sup> (Supplementary Table 3).

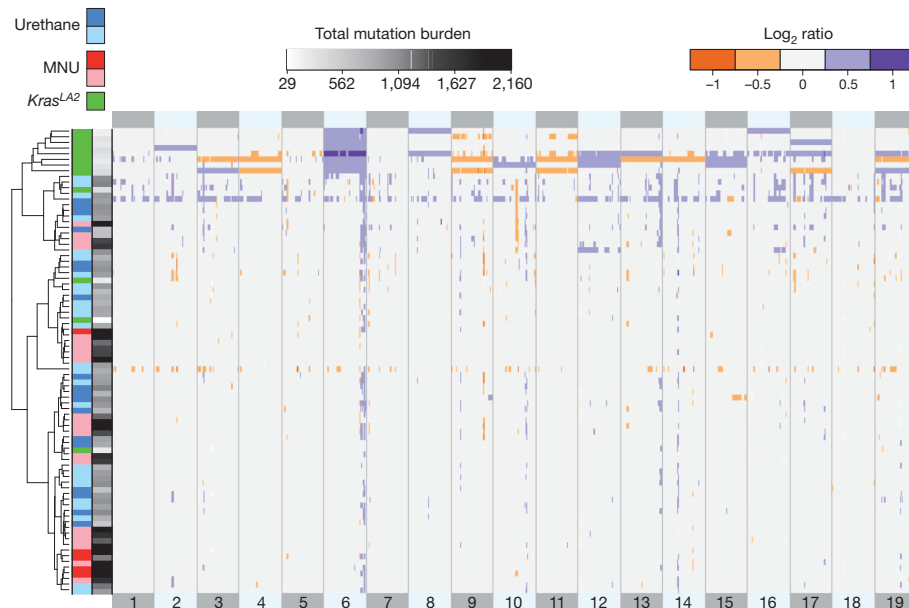
The observation that *Kras*<sup>LA2</sup> tumours have on average 15-fold fewer SNVs than MNU-induced tumours (Fig. 1b), despite sharing similar histology and the same *Kras* mutation, suggested that there are additional factors influencing tumorigenesis in these samples. Indeed, we found that copy number alterations (CNAs) are widespread in *Kras*<sup>LA2</sup> tumours (average = 3.25) but infrequent in carcinogen-induced tumours (average = 0.07), and hierarchical clustering by copy-number profile clearly segregated the carcinogen-induced and *Kras*<sup>LA2</sup> tumours into different groups (Fig. 2). Most *Kras*<sup>LA2</sup> tumours (9/12) showed amplification of *Kras*, mainly via gain of one copy of chromosome 6. These tumours

also carried common gains on chromosomes 2, 10, 12, 15 and 17, and deletions on chromosomes 4, 9, 11 and 17 (Extended Data Fig. 6), consistent with previously published array comparative genomic hybridization results from the *Kras*<sup>LA2</sup> model<sup>25</sup>. By contrast, carcinogen-induced tumours had very few CNAs and aneuploidies.

A summary of SNVs and CNAs involving driver genes reveals that all SNVs occurred in carcinogen-induced tumours and overwhelmingly showed the signature of the initiating carcinogen (Fig. 3). This suggests that carcinogen models produce tumours with a diversity of potential secondary driver SNVs, recapitulating in part the mutational heterogeneity seen in human cancer. One MNU-induced tumour harboured an E40K mutation in *Akt1*, generating a constitutively active oncoprotein<sup>26</sup>, and an early nonsense mutation in the tumour suppressor gene *Pax5*. Together with *Kras* G12D, this tumour had three functional mutations in cancer drivers, all MNU signature mutations probably induced in the same cell after MNU treatment. Although the *Kras*<sup>LA2</sup> tumours had no SNVs in established driver genes, some exhibited CNAs involving driver genes mutated in the carcinogen-induced tumours (Fig. 3). Further evidence for the role of CNAs in genetically engineered mouse models of cancer is provided by a recent report showing that mouse small-cell lung cancers induced by inactivation of *Trp53* and *Rb1* exhibit many CNAs, but a paucity of SNVs<sup>27</sup>. Similarly, mouse lung tumours induced by Cre activation of *Kras*<sup>LSL-G12D</sup> exhibit extremely few exome-wide SNVs (T. Jacks, personal communication). We conclude that carcinogen and genetic models show fundamental differences in patterns of genomic alterations, and that the requirement for CNAs may be abrogated by the high frequency of carcinogen-induced SNVs—a reciprocal relationship also seen in a recent analysis of TCGA sequencing of several thousand human tumours<sup>1</sup>.

To understand the processes operative in progression to adenocarcinoma, we performed WES on 9 FVB/N and 13 A/J strain urethane-induced, histologically confirmed lung adenocarcinomas (Extended Data Fig. 7a, b). The observed urethane signature A > G and A > T substitutions recapitulate the rates and patterns seen in the adenomas with remarkable fidelity (Extended Data Fig. 8), validating the utility of mouse carcinogen models to resolve complex mutational spectra. Further analysis revealed a significant increase in the CGN > A signature of genomic instability in both FVB/N and A/J adenocarcinomas (Fig. 4). This increase cannot be attributed to tumour age, as the FVB/N adenocarcinomas



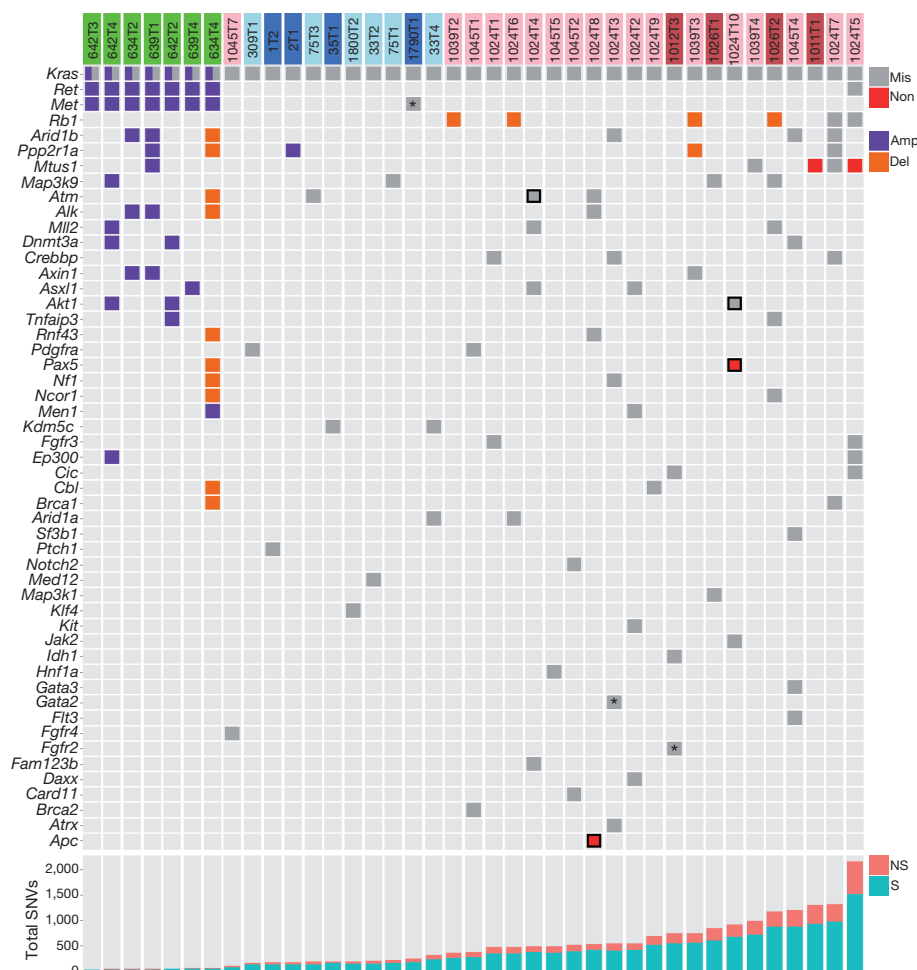


**Figure 2 | Distinct copy number profiles of genetically and chemically induced tumours.** Unsupervised, hierarchical clustering of  $\log_2$ -transformed read count ratios. *Kras*<sup>LA2</sup> tumours showed a significantly higher number of CNAs compared with carcinogen-induced tumours ( $P = 4.3 \times 10^{-10}$ , Wilcoxon rank-sum test). Chromosomes are aligned head to tail on the x axis, starting on the left. Samples are labelled by treatment and genotype, with *Kras*<sup>+/−</sup> samples appearing as light blue and light red. Sample SNV burden is displayed along the y axis in greyscale.

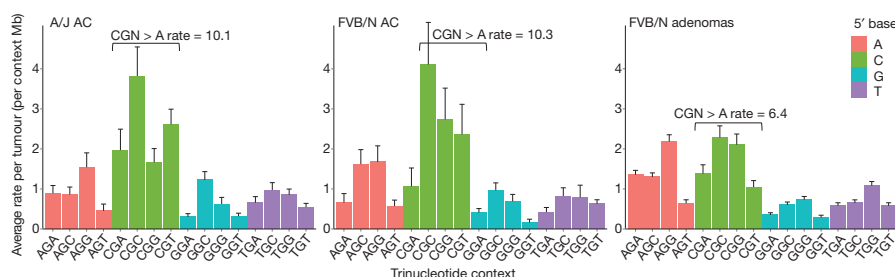
and adenomas were harvested at the same time point (20 weeks) after carcinogen treatment.

Most adenocarcinomas harboured Q61R mutations in *Kras* (Supplementary Table 4). Although urethane is known to induce *Kras* Q61L lung adenomas in A/J mice, adenocarcinomas from these animals harbour predominantly *Kras* Q61R mutations<sup>28</sup>. Eleven additional SNVs in driver genes were identified, as well as three SNVs in the reported

mouse lung adenoma suppressor gene *Fat4* (Supplementary Table 5)<sup>29</sup>. Compared with the urethane-induced adenomas, the adenocarcinomas are enriched for tumours with SNVs in high-likelihood driver genes other than *Kras* (Fisher  $P = 0.046$ ), as well as tumours harbouring CGN > A transitions in these genes (Fisher  $P = 0.034$ ). These data suggest that CGN > A transitions may play a part in the progression of adenomas to adenocarcinomas.



**Figure 3 | Consequential SNVs in high-likelihood driver genes only occur in carcinogen-induced tumours.** All missense (Mis) and nonsense (Non) SNVs, amplifications (Amp) and deletions (Del) in genes listed in Extended Data Table 2 are displayed. *Kras*<sup>LA2</sup> and urethane- and MNU-induced tumours are denoted at the top in green, blue and red, respectively, with lighter shading denoting *Kras*<sup>+/−</sup> genotype. SNVs with unequivocal evidence of consequence are bordered in black. All SNVs, excepting those marked with an asterisk, are concordant with the signature mutations of the inducing carcinogen. Bottom, total SNVs per tumour. NS, nonsynonymous; S, synonymous.



**Figure 4 | Adenocarcinomas show enrichment for a signature of genomic instability.** Analysis of G > A transitions in A/J and FVB/N adenocarcinomas (AC) reveals significant increases in CGN > A (NCG > T) transition rates compared to FVB/N adenomas ( $P = 0.00047$  and  $P = 0.0143$ , respectively, Wilcoxon rank-sum test), despite similar rates and patterns of other G > A transitions. Mutation counts per tumour were normalized to total length of sequenced trinucleotide contexts in each tumour and averaged. Error bars represent s.e.m.

A comparison of all validated carcinogen-induced mouse mutations with WES of human lung adenocarcinoma (TCGA LUAD,  $n = 230$ ) revealed substantial overlap in driver genes harbouring consequential mutations, both overall and in *KRAS*-mutant tumours (Supplementary Table 6). Some of the most frequently mutated genes in the mouse tumours (*Arid1b*, *Atm*, *Crebbp*, *Mll2* (also known as *Kmt2d*), *Rb1*) were also frequently mutated in the human tumours. Many of the mouse mutations occurred near mutations identified in TCGA LUAD, including functional mutations in *Akt1*, *Atm* and *Cbl* (Supplementary Table 7). In addition, the driver genes *ALK*, *APC*, *JAK2*, *MET* and *NF1*, commonly mutated in human NSCLC<sup>11</sup>, were mutated in the mouse tumours. Finally, an analysis of *MTUS1* mutations in TCGA LUAD revealed only consequential mutations (1.7%)—two missense mutations, and two frame-shift deletions—suggesting that loss-of-function mutations in *MTUS1* may be selected for in a subset of lung adenocarcinomas.

Genomic analysis of mouse tumours induced by a range of carcinogens may help reveal the relationships between environmental exposures and tumour architecture. Models that encompass heterogeneity in both genetic background and carcinogen exposure may also be useful for pre-clinical testing of cancer therapeutics, including immunotherapy, as the diversity of germline and somatic SNVs may recapitulate the variation in drug response and resistance observed in human clinical trials. Importantly, carcinogen models enable the production of tumours with a range of initiating *Ras* lesions, providing a valuable resource for interrogating the specificity and idiosyncrasies of these different mutations.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 December 2013; accepted 29 September 2014.

Published online 2 November 2014.

- Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nature Genet.* **45**, 1127–1133 (2013).
- Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).
- Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
- Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
- Johnson, L. *et al.* Somatic activation of the *K-ras* oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116 (2001).
- You, M., Candrian, U., Maronpot, R. R., Stoner, G. D. & Anderson, M. W. Activation of the *Ki-ras* protooncogene in spontaneously occurring and chemically induced lung tumours of the strain A mouse. *Proc. Natl Acad. Sci. USA* **86**, 3070–3074 (1989).
- Prior, I. A., Lewis, P. D. & Mattos, C. A comprehensive survey of *Ras* mutations in cancer. *Cancer Res.* **72**, 2457–2467 (2012).
- Jackson, E. L. *et al.* Analysis of lung tumour initiation and progression using conditional expression of oncogenic *K-ras*. *Genes Dev.* **15**, 3243–3248 (2001).
- Zhang, Z. *et al.* Wildtype *Kras2* can inhibit lung carcinogenesis in mice. *Nature Genet.* **29**, 25–33 (2001).
- To, M. D. *et al.* *Kras* regulatory elements and exon 4A determine mutation specificity in lung cancer. *Nature Genet.* **40**, 1240–1244 (2008).
- Govindan, R. *et al.* Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012).
- Forkert, P. G. Mechanisms of lung tumorigenesis by ethyl carbamate and vinyl carbamate. *Drug Metab. Rev.* **42**, 355–378 (2010).

- Kurowska, M., Labocha-Pawlowska, A., Grizda, D., Maluszynski, M. & Szarejko, I. Molecular analysis of point mutations in a barley genome exposed to MNU and gamma rays. *Mutat. Res.* **738–739**, 52–70 (2012).
- Pfeifer, G. P. Mutagenesis at methylated CpG sequences. *Curr. Top. Microbiol. Immunol.* **301**, 259–281 (2006).
- Welch, J. S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
- Drugan, J. K. *et al.* Ras interaction with two distinct binding domains in Raf-1 may be required for Ras transformation. *J. Biol. Chem.* **271**, 233–237 (1996).
- Spoerner, M., Herrmann, C., Vetter, I. R., Kalbitzer, H. R. & Wittinghofer, A. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl Acad. Sci. USA* **98**, 4944–4949 (2001).
- Fawad, S. *et al.* Targeted genetic dependency screen facilitates identification of actionable mutations in FGFR4, MAP3K9, and PAK5 in lung cancer. *Proc. Natl Acad. Sci. USA* **110**, 12426–12431 (2013).
- Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
- Di Benedetto, M. *et al.* Mutation analysis of the 8p22 candidate tumour suppressor gene *ATIP/MTUS1* in hepatocellular carcinoma. *Mol. Cell. Endocrinol.* **252**, 207–215 (2006).
- Frank, B. *et al.* Copy number variant in the candidate tumour suppressor gene *MTUS1* and familial breast cancer risk. *Carcinogenesis* **28**, 1442–1445 (2007).
- Zuern, C. *et al.* Down-regulation of *MTUS1* in human colon tumours. *Oncol. Rep.* **23**, 183–189 (2010).
- Xiao, J. *et al.* Reduced expression of *MTUS1* mRNA is correlated with poor prognosis in bladder cancer. *Oncol. Lett.* **4**, 113–118 (2012).
- Shedden, K. *et al.* Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nature Med.* **14**, 822–827 (2008).
- To, M. D. *et al.* Progressive genomic instability in the *FVB/Kras<sup>L2</sup>* mouse model of lung cancer. *Mol. Cancer Res.* **9**, 1339–1345 (2011).
- Ahmed, N. N., Grimes, H. L., Bellacosa, A., Chan, T. O. & Tsichlis, P. N. Transduction of interleukin-2 antiapoptotic and proliferative signals via Akt protein kinase. *Proc. Natl Acad. Sci. USA* **94**, 3627–3632 (1997).
- McFadden, D. G. *et al.* Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell* **156**, 1298–1311 (2014).
- Nuzum, E. O., Malkinson, A. M. & Beer, D. G. Specific *Ki-ras* codon 61 mutations may determine the development of urethan-induced mouse lung adenomas or adenocarcinomas. *Mol. Carcinog.* **3**, 287–295 (1990).
- Berndt, A. *et al.* Identification of *Fat4* and *Tsc22d1* as novel candidate genes for spontaneous pulmonary adenomas. *Cancer Res.* **71**, 5779–5791 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by National Cancer Institute (NCI) grants R01 CA111834, U01 CA84244, U01 CA141455 and U01 CA176287 (to A.B.), and partly funded by the Bonnie Addario Foundation. P.M.K.W. was supported by the National Institutes of Health (NIH) training grant T32 GM007175 and a National Science Foundation GRFP award, and is currently supported by an NCI F31 NRSA award. K.D.H. was supported by the NIH training grant T32 GM007175, and is currently supported by an NCI F31 NRSA award. D.J.A. is supported by Cancer Research UK and the Wellcome Trust. We are appreciative of help and comments from our colleagues in refining this study and manuscript. We would also like to thank S. Busch for assistance with animal studies, and S. Green, T. Yuan and M. McMahon for providing the K493.1 cell line.

**Author Contributions** P.M.K.W., K.D.H., M.D.T., D.J.A. and A.B. contributed to the overall study design. P.M.K.W. carried out most of the experiments, with help from M.D.T. R.D. was responsible for all of the animal studies. Sequencing and Sequenom were performed at the Sanger Institute under the supervision of D.J.A., and data processing was carried out by K.D.H., M.R., A.G.R. and T.M.K. SNV and CNA calling were carried out by K.D.H. Data analysis was carried out primarily by P.M.K.W. and K.D.H., with help from E.F. and D.A.Q. K.-Y.J. made histological assessments of all tumours. Adenomas and adenocarcinomas from the A/J mice were provided by C.J.K. and K.E.G. The manuscript was written primarily by P.M.K.W. and A.B., with contributions from the other authors.

**Author Information** The raw .bam files have been deposited in the European Nucleotide Archive under accession number ERP001454. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.B. (abalmain@cc.ucsf.edu).

## METHODS

**Mouse strains and tumour induction.** *Kras*<sup>LA2</sup> and *Kras*<sup>LSL-G12D</sup> alleles, originally on a C57BL6/129/SvJae background, were backcrossed onto the FVB/N genetic background for more than 20 generations. Mice were treated with urethane (1 g kg<sup>-1</sup>) or MNU (50 mg kg<sup>-1</sup>) dissolved in PBS by intraperitoneal injection at ~7–12 weeks of age. Lung tumours from mice induced with carcinogen were harvested at ~20 weeks after injection, or ~32 weeks in the A/J animals, while spontaneous lung tumours were collected from *Kras*<sup>LA2</sup> mice at ~9 months of age. For the urethane-induced adenomas, 18 tumours from seven wild-type animals and 26 tumours from nine *Kras*<sup>LSL-G12D</sup> animals were collected. For the MNU treatment group, 5 tumours from four wild-type animals and 21 tumours from three *Kras*<sup>LSL-G12D</sup> animals were collected. A total of 12 tumours were collected from four *Kras*<sup>LA2</sup> animals. Eight histologically confirmed adenocarcinomas were collected from four FVB/N *Kras*<sup>LSL-G12D</sup> animals, and one from a wild-type FVB/N animal. Thirteen tumours, including ten histologically confirmed adenocarcinomas, were collected from seven wild-type A/J animals. *Kras*<sup>LSL-G12D</sup> is a latent G12D allele that is inactive in the absence of Cre-recombinase. Importantly, lungs from *Kras*<sup>LSL-G12D</sup> heterozygous mice were shown to have an approximately twofold reduction of *Kras* messenger RNA transcript and protein compared with wild-type littermates<sup>30</sup>. Furthermore, these mice had more and larger lung tumours than wild-type mice after carcinogen treatment<sup>30</sup>, similar to the results seen for animals heterozygous for the original *Kras* null allele<sup>31</sup>.

No animals or tumours were excluded from the analysis. Tumours were collected from male and female mice, and no sex differences were observed. No formal randomization was performed, and all analyses were performed against the entire set of data in an unbiased manner. All animal experiments were approved by the University of California San Francisco Laboratory Animal Resource Center.

**DNA isolation and sequencing.** Formalin-fixed or flash-frozen tumours free of visible normal tissue were digested overnight in proteinase K (Bioline) and phenol/chloroform purified using 5 PRIME Phase Lock Gel Heavy Tubes (Fisher Scientific). Integrity of genomic DNA was assessed by electrophoresis on 1% agarose gels, and concentration was determined by nanodrop spectrophotometry and PicoGreen (Invitrogen). Exome enrichment and sequencing genomic libraries were prepared using the Illumina Paired End Sample Prep Kit following the manufacturer's instructions. Enrichment was performed as described previously<sup>32</sup> using the Agilent SureSelect Mouse All Exon kit following the manufacturer's recommended protocol. Each exome was sequenced using a 76 bp paired-end protocol on the Illumina platform (GAII or HiSeq2000).

**Sequence alignment, processing and quality control.** Tumour .bam files were aligned to the GRCh38/mm10 version of the *Mus musculus* genome using BWA (version 0.5.9)<sup>33</sup>. After alignment, duplicates were marked and mate information was fixed using Picard (version 1.80; <http://broadinstitute.github.io/picard/>). We then recalibrated base quality score and realigned reads around indels using GATK (version 2.2-15)<sup>34</sup>. Finally, alignment and coverage metrics were collected using Picard. We sequenced an average of 75 million unique on-target reads per tumour. Targeted bases were sequenced to a mean depth of 72, and more than 88% of targeted bases were sequenced to 20× coverage or greater. There were no significant differences in depth of coverage or proportion of regions covered to 20× between tumour induction groups.

**Identification of SNVs and annotation.** SNVs were identified using the somatic variant detection program, MuTect (version 1.1.4)<sup>35</sup>. Tumours were called against DNA taken from normal tail isolated from two wild-type FVB/N control samples. GRCh38/mm10 served as the reference during calling. Each set of variants was then subset to those variants that passed MuTect filters and had a minimum read depth of 12. The intersection of both callsets was then filtered for known variants from the database of mouse variation available at <http://ftp-mouse.sanger.ac.uk> (release 1303, mmp.v3). Variants found only in *Mus spretus*, *Mus castaneus* or *Mus musculus musculus* were not used for filtration. All samples were also filtered for variants observed in a panel of six controls. These comprised the two wild-type samples used for variant calling, two *Kras*<sup>LA2</sup> mice, and two *Kras*<sup>LSL-G12D</sup> heterozygous mice. These mice were then called for variants using FreeBayes (version 0.9.8; <http://arxiv.org/abs/1207.3907>), UnifiedGenotyper (version 2.2-15)<sup>34</sup> and mpileup (version 0.1.18)<sup>36</sup>. Variants from each caller were then filtered for sites with a minimum quality of 50 and minimum depth of 10. Variants called by a minimum of two callers were used to filter variants. Surviving variants were annotated using Annovar (downloaded on 5 September 2013)<sup>37</sup>. A final level of filtration was performed on variants that showed clear clustering by mouse, which were called SNPs and discarded. In *Kras*<sup>LSL-G12D</sup> mice, MNU-induced G12D mutation of the wild-type allele was clearly distinguished from latent G12D on the *Kras*<sup>LSL-G12D</sup> allele by observation of a nearby SNP, unique to the *Kras*<sup>LSL-G12D</sup> allele, in the exome-sequencing reads as well as Sanger sequencing.

**Mutation spectra analysis.** SNVs in all tumours were annotated by the 96 possible trinucleotide context substitutions (6 types of substitutions × 4 possible flanking

5' bases × 4 possible flanking 3' bases) and summed in each tumour, creating a matrix of 82 tumours × 96 substitutions. For hierarchical clustering, these counts were converted to per tumour proportions and clustered by Euclidean distance and similarity computed by nearest neighbour in R. For heatmaps in Fig. 1c and Extended Data Fig. 1, substitution counts were log<sub>10</sub> normalized, column scaled and centred on 0. Mutation spectra barplots were created by dividing each totalled type of substitution in each tumour by the total number of successfully sequenced contexts (defined as ≥10× coverage) in that tumour corresponding to each substitution, retrieved from mpileup of the .bams in samtools. The resulting per tumour substitution rates were then averaged across all tumours in the respective treatment groups.

**Prioritization of high-likelihood driver genes.** We explored a recently published gene prioritization approach that specifically addresses the phenomenon of spurious enrichment of longer genes by adjusting for gene expression and replication timing<sup>4</sup>. However, given the scarcity of recurrent variants in our data set limiting the utility of this approach, we decided to prioritize variants that occurred in genes described by Vogelstein and colleagues as known to harbour bona fide driver mutations in cancer<sup>19</sup>, as well as the recently identified lung cancer driver genes *Fgfr4*, *Map3k9* and *Pak5* (ref. 18). In particular, Vogelstein *et al.* described a stringent list of 125 driver genes harbouring subtle mutations based on the criteria that >20% recorded mutations in oncogenes must be recurrent and missense, and >20% recorded mutations in tumour suppressors must be inactivating. *Mtus1* was chosen for further investigation owing to the recurrence of missense and nonsense mutations. Variants were compared to known human somatic mutations as available via the COSMIC database<sup>38</sup>. Briefly, the mouse and human sequences for homologous proteins were pairwise aligned using Clustal Omega<sup>39</sup> and the human protein position homologous to the mouse mutation was used to query COSMIC for known missense and nonsense mutations at or surrounding this peptide position. Local conservation was determined after sequence alignment using a ± 10 amino acid residue window surrounding the substituted amino acid.

**Validation of SNVs.** SNVs were validated by either Sequenom MassARRAY or conventional Sanger sequencing. SNVs were called validated if they were detected in the tumour but not in matched normal DNA. A subset of SNVs that failed both methods for technical reasons was called validated if individual inspection of the aligned reads in tumours and controls strongly supported validity, as performed in previous studies<sup>40</sup>. The method used for validation of SNVs in driver genes is noted in Extended Data Table 2 and Supplementary Table 5. Altogether, validation was attempted on 401 SNVs from the adenomas. A total of 11 failed for technical reasons, and 13 were inconclusive. A total of 17 variants were validated by visual inspection, representing 4.2% of the 401 variants tested. SNVs tested by Sequenom were called inconclusive if the SNV was observed in the tumour but failed in the control, or the SNV was observed in the tumour and not the matched normal control, but was observed in control tissue from another mouse. SNVs tested by visual inspection were called inconclusive if inspection suggested somatic origin, but total variant reads were less than 10. The overall validation rate (excluding inconclusive SNVs) was 87%. The Sequenom validation rate alone was 86%. The vast majority of *Kras* mutations were validated by Sanger sequencing, although a small subset went undetected by this method (Supplementary Table 1) despite confirmation by manual inspection of the alignments, suggesting a higher sensitivity afforded by WES. These patterns confirm previous results on carcinogen-specific mutations in *Kras*<sup>6,9,10</sup>. Sanger sequencing validation was attempted on 20 randomly selected CGN > A transitions as well as 3 CGN > A transitions in driver genes in the adenocarcinoma samples, 15 of which passed (Supplementary Table 8). Alignments were visually inspected for the remaining eight, all of which supported somatic origin, but only one of which had enough variant reads (≥10) to pass. Interestingly, the inconclusive variants and the majority of the validated variants had very low variant read fractions, supporting a hypothesis that the CGN > A mutations were acquired during progression and are represented in subclonal tumour fractions.

**Assessment of copy number from read depth.** Copy number was estimated from sequencing data using FREEC (version v6.4; <http://bioinfo-out.curie.fr/projects/freec/>). Read depth was compared between tumour and control samples to estimate copy number in 8 kb windows, and subsequently segmented via a LASSO-based algorithm<sup>41</sup>. FREEC was run with the following parameters: window size, 8 kb; step size, 2.5 kb; contaminationAdjustment = TRUE; noisyData = TRUE; BAF calculation activated. 2.5 kb windows were then aggregated into 15 kb bins by taking the median ratio for all covered windows. Each tumour was profiled against the two wild-type controls used for variant calling. Aggregate profiles were generated for each tumour by the following rules: if either ratio was approximately neutral, the region was considered neutral; if both ratios were aberrant with the same directionality, the more conservative ratio was used; if both ratios were aberrant with different directionality, the region was discarded. Resulting merged ratios were then inspected for high missing rates and low variance, which were then omitted. Additionally, several small regions with evidence of technical artefacts resulting in extremely consistent aberration rates (greater than 50% of samples) across all treatment



groups were manually excluded. Particularly, these regions were manually inspected for the existence of large gene families that could account for misalignments and result in spurious aneuploidies. Short spans on chromosomes 1, 4, 6 and 12 were discarded as artefacts.

**Histological classification.** A small piece of each tumour was collected and paraffin embedded for pathology, sectioned to 6 µm and haematoxylin and eosin (H&E) stained. Histological architecture was classified as either papillary, solid, or mixed papillary and solid. Solid was defined as histology with marked lack of papillary structure, yet more structure than traditionally solid lung adenocarcinomas in humans. Adenocarcinomas were called based on large size and the presence of the following cytological criteria: tumour cell crowding, scattered mitotic figures, nuclear atypia (enlargement and moderate pleomorphism), nuclear membrane irregularity, and prominent nucleoli. All histology was called by a lung pathologist blinded to the study groups and conditions.

**Cell culture, *Mtus1* knockdown and MTT assay.** The mouse lung cancer cell line K493.1, which harbours a *Kras* G12D mutation, was grown in DMEM supplemented with 10% fetal bovine serum (Atlas Biologicals). *Mtus1* was knocked down using 50 nM ON-TARGETplus SMARTpool siRNA (Dharmacon) containing multiple pooled siRNAs targeting all isoform transcripts of mouse *Mtus1* (catalogue number L-065229-01). Transfection of siRNA was performed at ~20–40% cell confluence using Lipofectamine-2000 (Invitrogen). In parallel, cells were transfected with control ON-TARGETplus Non-Targeting Pool siRNA (catalogue number D-001810-10). RNA was harvested from cells at day 3 after transfection using Trizol reagent (Invitrogen), DNA was removed using the TURBO DNA-free kit (Ambion), and complementary DNA was synthesized from 500 ng RNA using the Superscript III First-strand Synthesis kit (Invitrogen). qRT-PCR was performed on cDNA using TaqMan Assays-on-Demand (Applied Biosystems) against mouse *Mtus1* (Mm00628662\_m1 *Mtus1*) and  $\beta$ -actin on the 7900HT Fast Real-Time PCR System (Applied Biosystems). Reactions were performed in quadruplicate, and levels of *Mtus1* were normalized to  $\beta$ -actin. Cell proliferation was assayed in 96-well plates (six replicate wells per group) at days 1, 2 and 3 after siRNA transfection using MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (Invitrogen). Formazan crystals were re-suspended in dimethylsulphoxide (DMSO), and absorbance was read at 540 nm. Four independent experiments were performed, and a significant increase in absorbance (Wilcoxon rank-sum test) was always seen in the *Mtus1* knockdown compared with control siRNA cells at day 3. One representative experiment is shown in Extended Data Fig. 5b. All protocols were performed following the manufacturer's instructions.

**Survival analyses in human lung cancer data sets.** The TCGA LUAD and human lung squamous cell carcinoma (LUSC) data sets were downloaded from the UCSC Cancer Genomics Browser (<https://genome-cancer.ucsc.edu>). Illumina HiSeq 2000 RNA Sequencing expression data was used for analyses of gene expression with overall survival. A validation data set for *MTUS1* expression in lung adenocarcinoma<sup>24</sup> was downloaded from <https://caintegrator.nci.nih.gov/caintegrator/>. Analysis of *MTUS1* expression and survival was also repeated in a second squamous cell carcinoma (SCC) data set<sup>42</sup>, which was downloaded from the UCSC Cancer Genomics Browser. No association between *MTUS1* expression and survival was seen in the SCC data sets (Supplementary Table 3), suggesting that *MTUS1* expression may only have prognostic significance in certain types of lung cancer such as NSCLC. For all survival analyses, clinical covariates of sex, age, cigarette pack years smoked and stage were included, except in one data set<sup>24</sup> in which cigarette pack years smoked was not available. Cox regression was performed in R with gene expression as a continuous variable. High and low expression groups were split about median expression values for plotting Kaplan–Meier curves.

**Human versus mouse mutation comparison.** Genes included in this comparison were limited to known driver genes (see earlier) harbouring mutations in the

carcinogen-induced mouse tumours. The TCGA LUAD WES .vcf was downloaded from the UCSC Cancer Genomics Browser. Only functional SNVs and indels were included. Validated functional SNVs from carcinogen-induced mouse adenomas and adenocarcinomas, and CNAs from the carcinogen-induced adenomas, were used in the comparison. Inclusion of mouse CNAs (six total) made little difference overall, but were included to emphasize recurrent mutation of *Rb1* in the mouse tumours, which had four deletions and two missense SNVs.

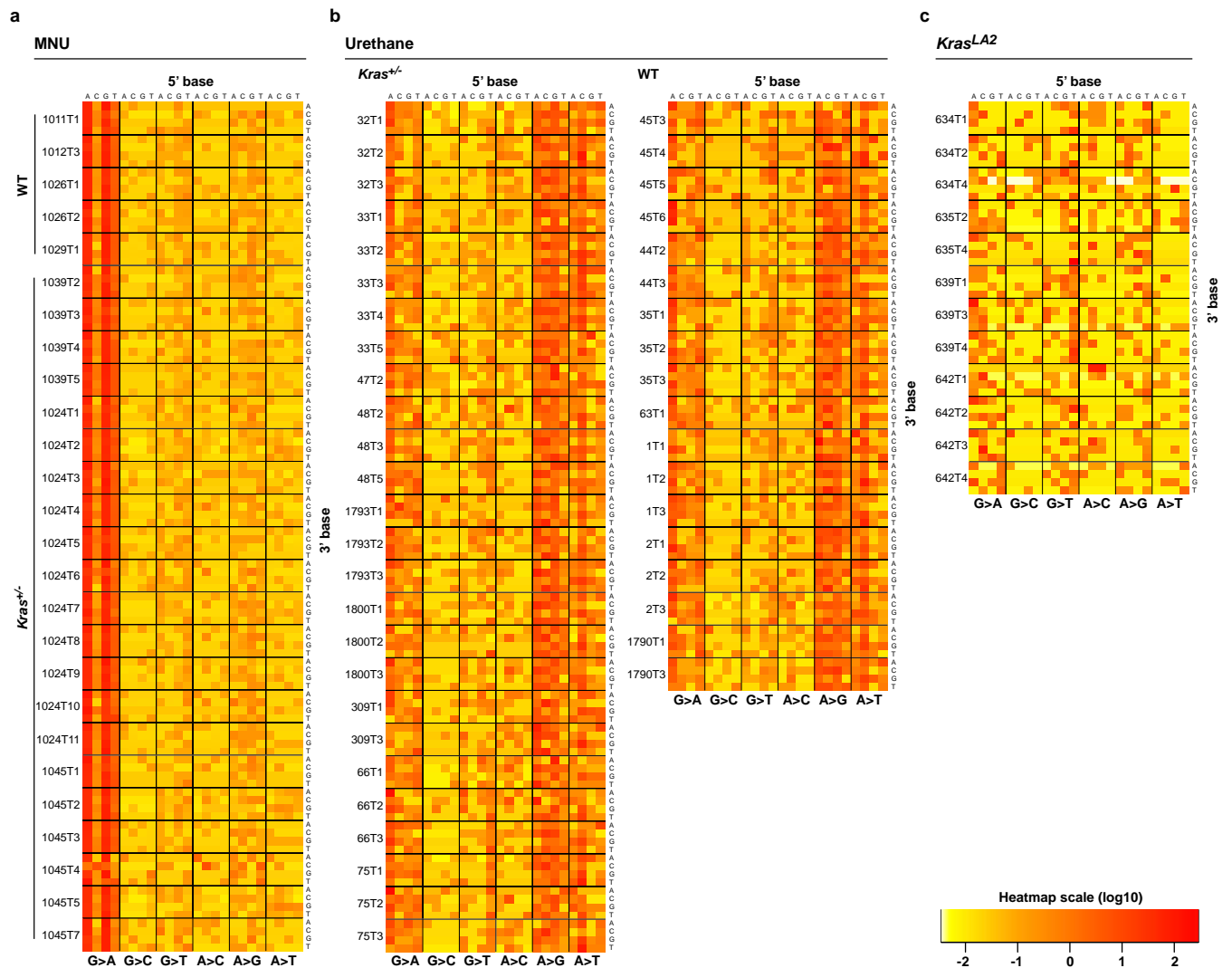
**Generation of plots.** All plots were created using the statistical computing language R (<http://www.R-project.org/>). Heatmaps were generated using the heatmap.2 function in the gplots package (<http://cran.r-project.org/web/packages/gplots/index.html>), Kaplan–Meier curves were generated using the survival package (<http://CRAN.R-project.org/package=survival>), and all other plots were made using the ggplot2 package (<http://ggplot2.org/>).

**Statistical analyses.** The nonparametric Wilcoxon rank-sum test (Mann–Whitney U test) was used in Figs 1, 2, 4, and Extended Data Figs 2, 4, 5 for testing the alternative hypothesis that two populations of values differ against the null hypothesis that they are the same. This test was chosen owing to its efficiency in handling both normal and non-normal distributions. The Fisher exact test was used in the text and in Extended Data Figs 3 and 4 to compare count data between groups, and was chosen for its robust ability to handle high and low ranges of count data. Where appropriate, *P* values were adjusted for multiple tests using the Holm's correction for multiple comparisons. Survival analysis in Extended Data Fig. 5 is explained earlier. All data were visualized in R using summary statistics and basic plotting functions before statistical testing, and variance was comparable in all cases in which the Wilcoxon rank-sum test was used. All assumptions of statistical tests were met.

**Data deposition.** Aside from the raw .bam files (European Nucleotide Archive (ENA) accession ERP001454), a sample ID key with study names and ENA names is provided in Supplementary Information (ExomeLungTumorIDs\_Key.txt). Variant call format files of SNVs used in analyses in the paper are provided in Supplementary Information (Adenomas\_variants.txt, Adenocarcinomas\_variants.txt).

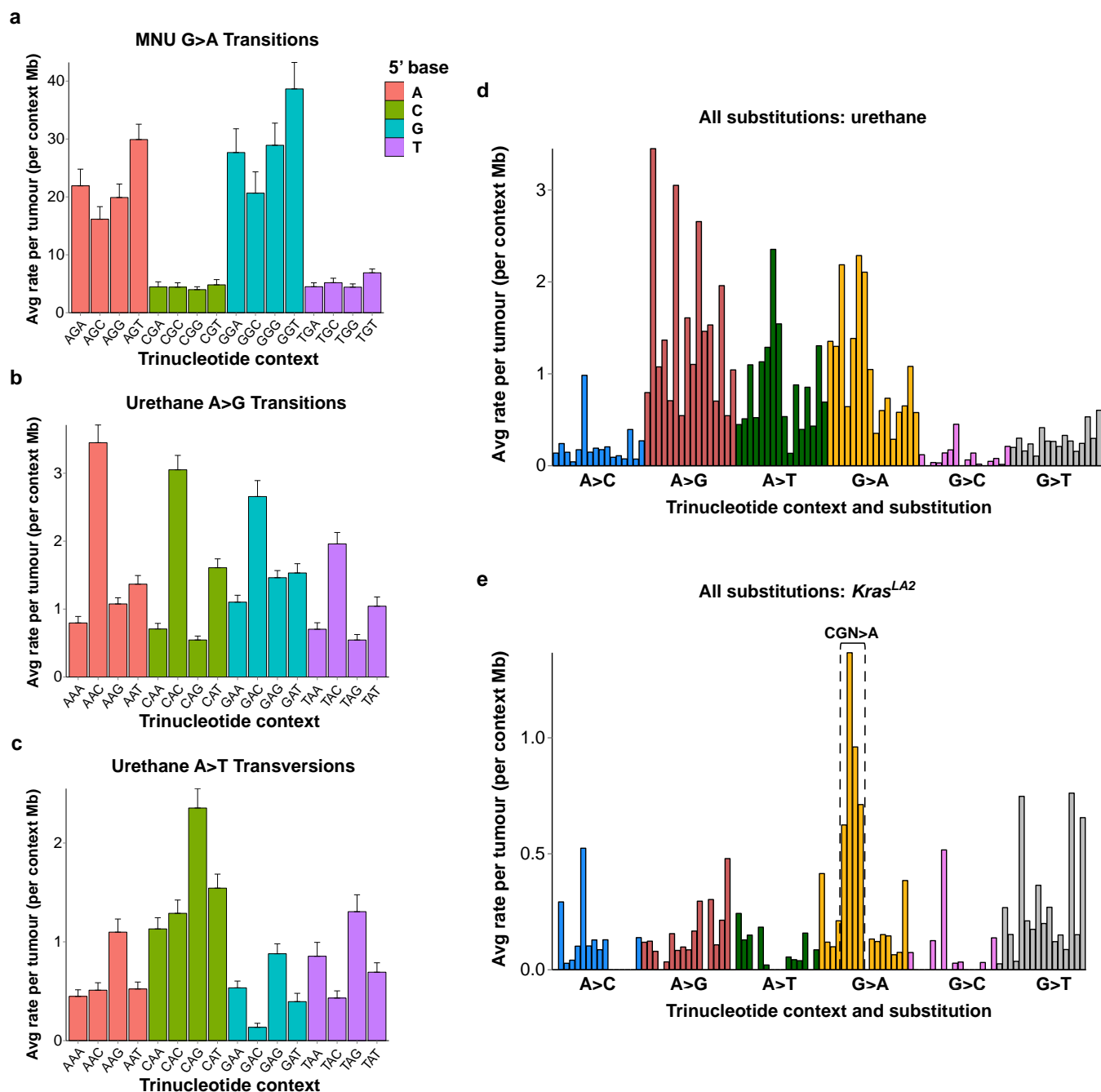
30. To, M. D., Rosario, R. D., Westcott, P. M., Banta, K. L. & Balmain, A. Interactions between wild-type and mutant *Ras* genes in lung and skin carcinogenesis. *Oncogene* **32**, 4028–4033 (2013).
31. Johnson, L. *et al.* *K-ras* is an essential gene in the mouse with partial functional overlap with *N-ras*. *Genes Dev.* **11**, 2468–2481 (1997).
32. Varela, I. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature* **469**, 539–542 (2011).
33. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
34. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43**, 491–498 (2011).
35. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol.* **31**, 213–219 (2013).
36. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
37. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
38. Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2011).
39. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
40. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
41. Boeva, V. *et al.* Control-free calling of copy number alterations in deep-sequencing data using GC-content normalization. *Bioinformatics* **27**, 268–269 (2011).
42. Raponi, M. *et al.* Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res.* **66**, 7466–7472 (2006).





**Extended Data Figure 1 | Distinct and consistent mutation spectra across tumours from carcinogen and genetic models.** a–c, Stacked heatmaps displaying the mutation spectra of all MNU-induced (a), urethane-induced (b), and *Kras*<sup>LA2</sup> tumours (c), shown as normalized frequencies of all 96

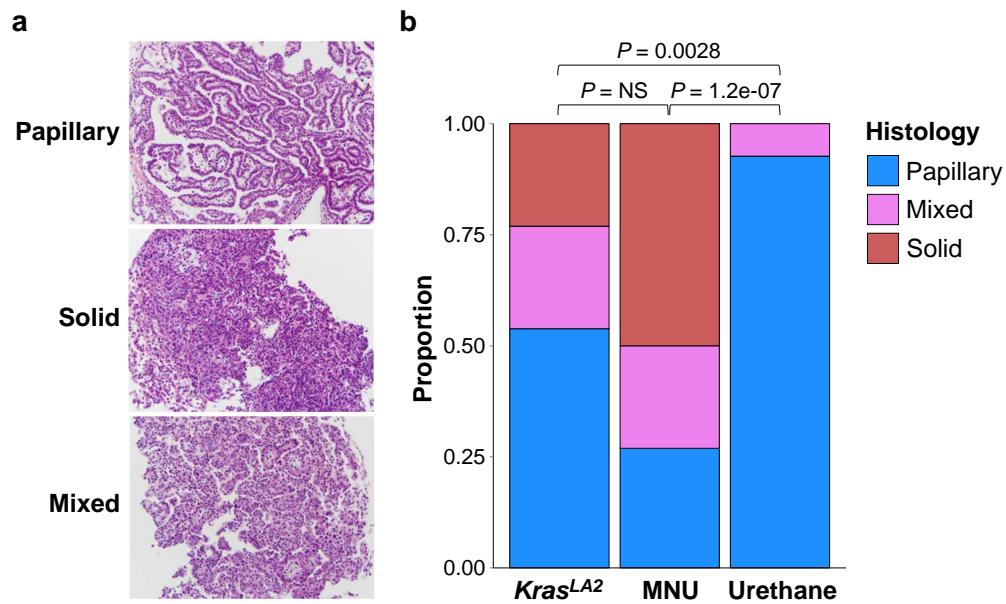
possible substitutions. Substitutions are shown below each heatmap, with 5'- and 3'-flanking base context displayed on the top and right, respectively. Tumour identifier is shown to the left of each heatmap.



# Extended Data Figure 2 | Highly specific mutation signatures.

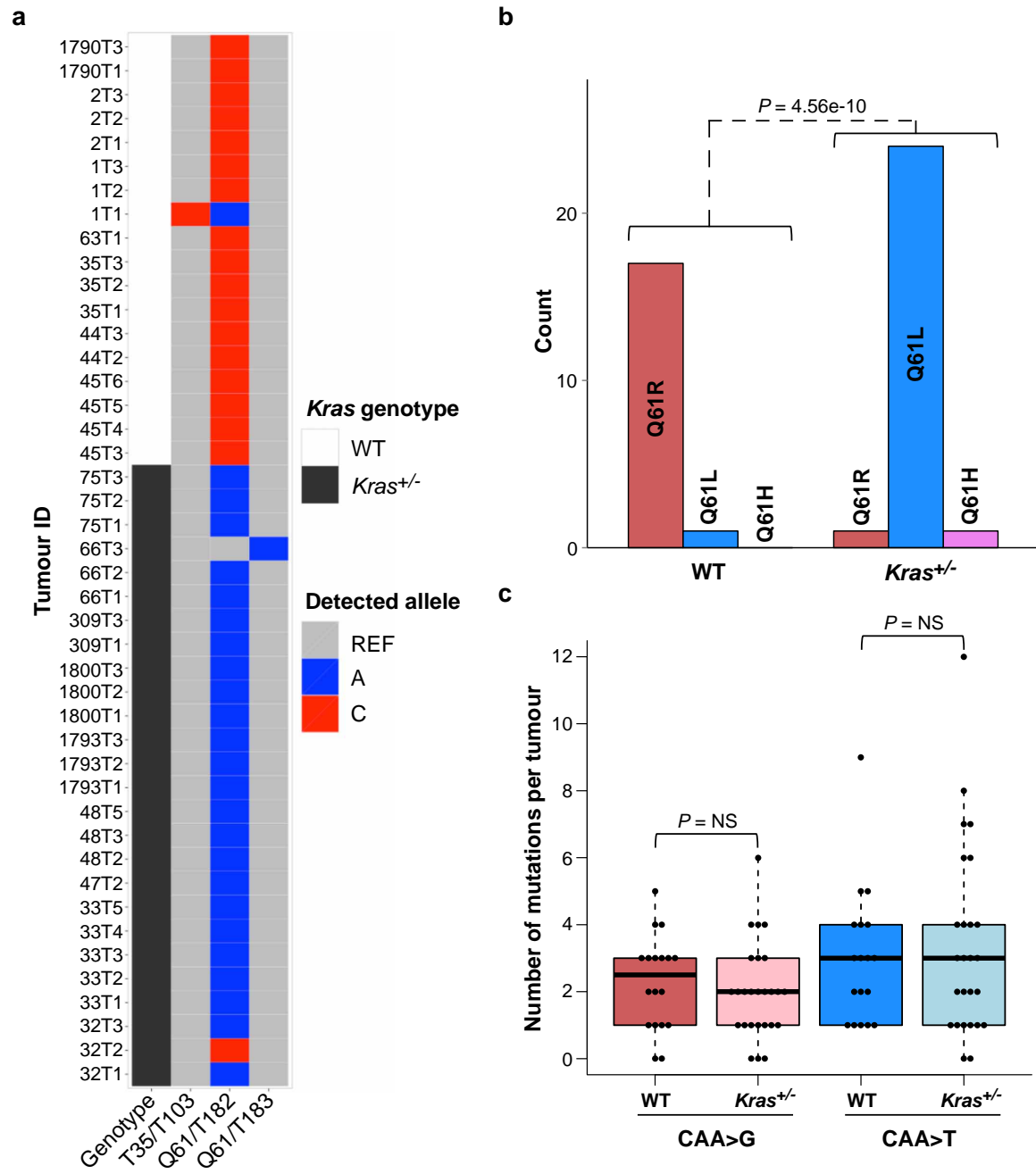
**a**, Breakdown of G > A transitions in MNU-induced tumours. 5'-flanking purine versus pyrimidine G > A substitutions, and 3'-flanking thymidine versus all other G > A substitutions, are highly significant ( $P < 0.0003$ , Wilcoxon rank-sum test). **b**, **c**, Breakdowns of A > G transitions (**b**) and

A > T transversions (**c**) in urethane-induced tumours. **d**, **e**, All 96 substitutions in urethane-induced (**d**) and *Kras*<sup>LA2</sup> tumours (**e**). **e**, The CGN > A (NCG > T) signature mutations of genomic instability are denoted. Mutation counts per tumour were normalized to total length of sequenced trinucleotide contexts in each tumour and averaged. Error bars represent s.e.m.



**Extended Data Figure 3 | *Kras* G12D mutation induces tumours with different histologies compared with codon 61 mutants.** **a**, Representative papillary, solid, and mixed tumour histologies ( $\times 200$  magnification). **b**, Breakdown of different histologies in each treatment group. Histologies

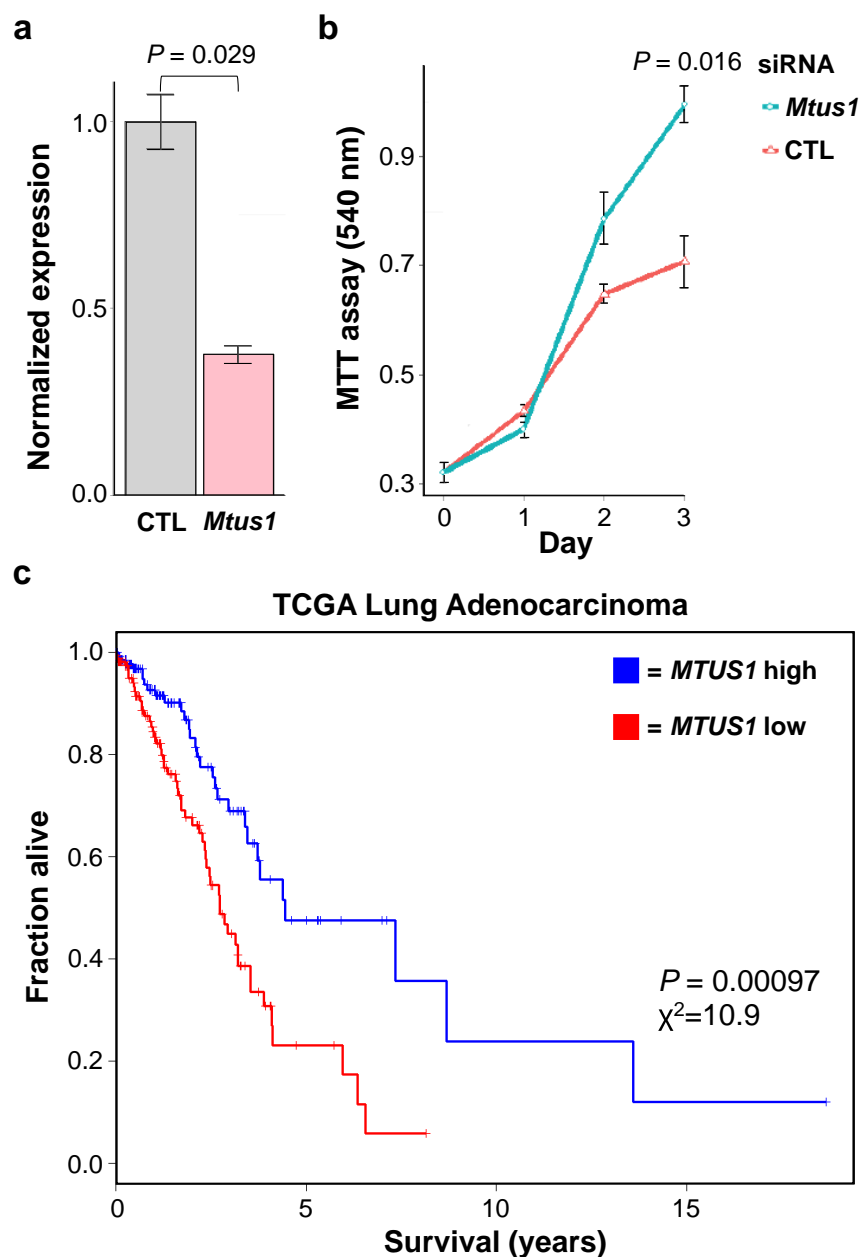
from  $Kras^{LA2}$  and MNU groups were significantly different compared with those from urethane, but there was no significant difference between  $Kras^{LA2}$  and MNU (Fisher exact test, Holm's correction for multiple comparisons).



**Extended Data Figure 4 | Germline *Kras* genotype influences mutation specificity in urethane-induced tumours.** **a**, *Kras* mutant alleles for urethane tumours are plotted as coloured squares for all three oncogenic alleles detected in these tumours. *Kras* genotype is indicated as either white (wild type (WT)) or black (heterozygous) squares. **b**, Highly significant switch in *Kras* codon 61

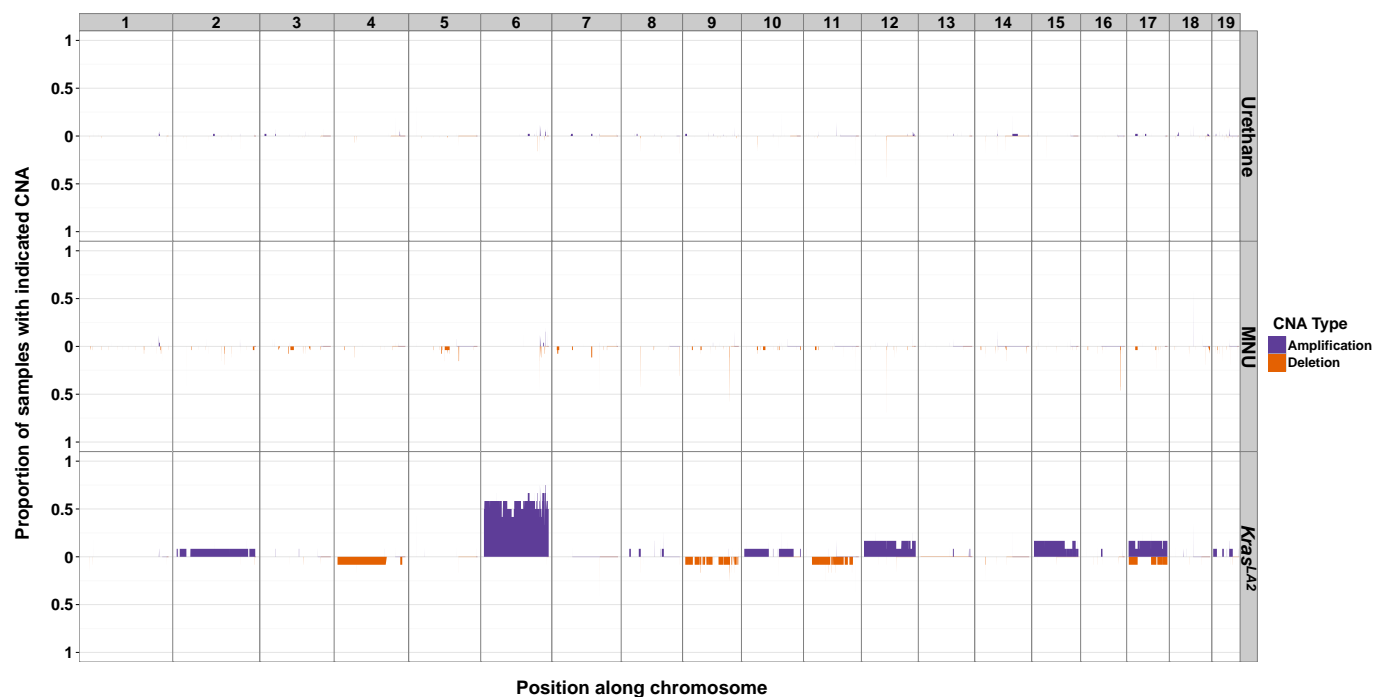
mutations between tumours from wild-type mice and *Kras*<sup>+/-</sup> mice (Fisher exact test). **c**, No significant difference was seen between the exome-wide rates of the substitutions underlying *Kras* Q61R (CAA > G) and Q61L (CAA > T) mutations between tumours from wild-type and *Kras*<sup>+/-</sup> mice (Wilcoxon rank-sum test). NS, not significant.



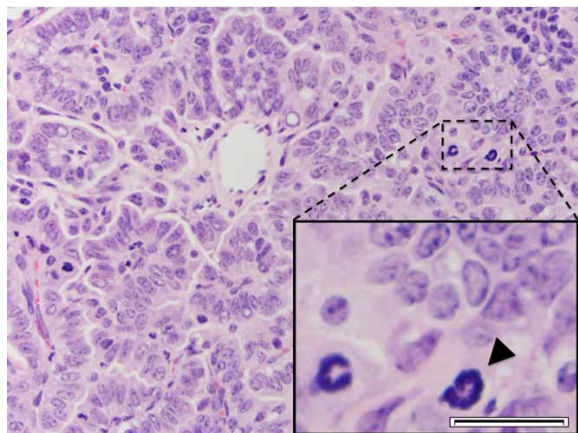
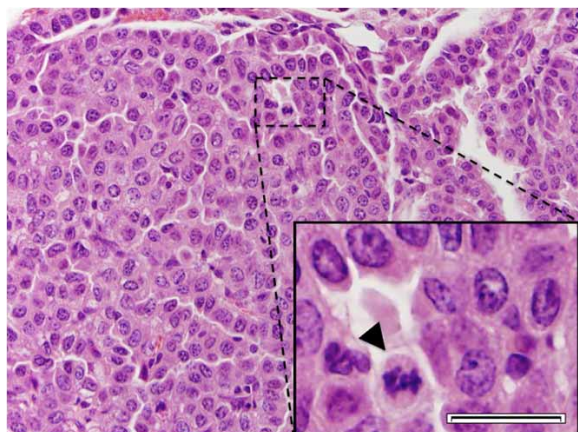


**Extended Data Figure 5 | *MTUS1* is a tumour suppressor in mouse and human lung cancer.** **a**, Polymerase chain reaction with quantitative reverse transcription (qRT-PCR) quantification of short interfering RNA (siRNA) knockdown of *Mtus1* in a *Kras* G12D mouse lung cancer cell line (K493.1) (Wilcoxon rank-sum test). CTL, control. **b**, MTT assay shows increased growth after *Mtus1* knockdown (Wilcoxon rank-sum test). Four independent

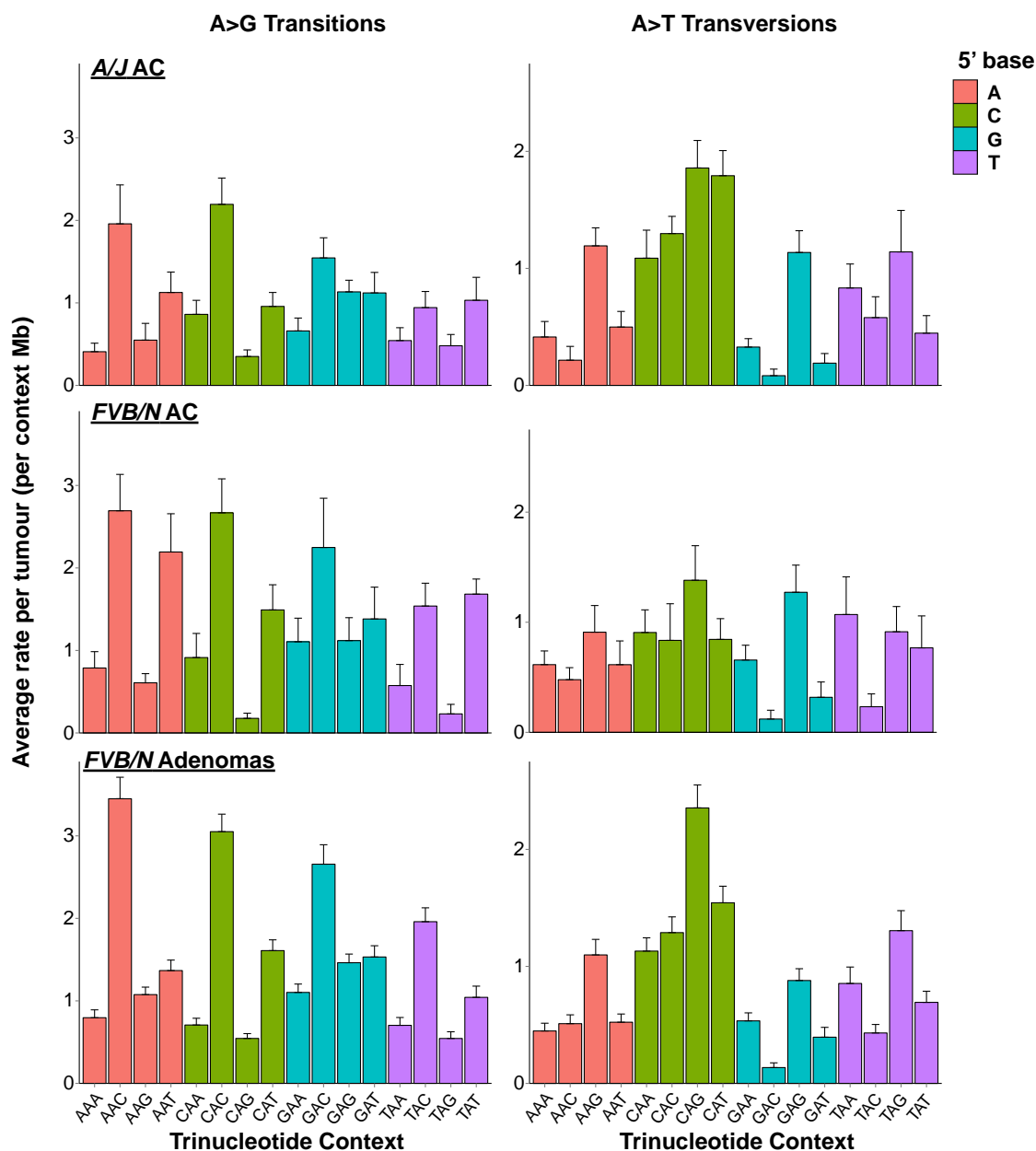
trials were performed and growth was significantly increased by day 3 after knockdown in each experiment. One representative trial is shown. **c**, *MTUS1* expression is significantly associated with overall survival in human lung adenocarcinoma ( $P = 0.00097$ ,  $\chi^2 = 10.9$ ). Analysis was performed using the clinical covariates gender, age, pack years smoked, and stage.



**Extended Data Figure 6 | Proportion of tumours with CNAs in each treatment group.** Amplifications and deletions were defined as regions with a  $\log_2$  ratio greater than 0.5 or less than  $-0.5$ , respectively. Chromosomes are arranged on the  $x$  axis in a head-to-tail formation.

**a****b**

**Extended Data Figure 7 | Histological confirmation of lung adenocarcinomas.** **a, b**, Representative histologies ( $\times 400$  magnification) of A/J (**a**) and FVB/N (**b**) adenocarcinomas. Zoom insets show tumour cell crowding and scattered mitotic figures (black arrowheads), nuclear atypia including enlargement and moderate pleomorphism, nuclear membrane irregularity, and prominent nucleoli. Scale bars, 20  $\mu\text{m}$ .



**Extended Data Figure 8 | Comparison of urethane signature mutations in adenomas and adenocarcinomas.** Urethane A > G transitions (left) and A > T transversions (right) are shown in A/J adenocarcinomas, FVB/N

adenocarcinomas and FVB/N adenomas. Mutation counts per tumour were normalized to total length of sequenced trinucleotide contexts in each tumour and averaged. Error bars represent s.e.m.



Extended Data Table 1 | Treatment groups and lung tumours for WES

Treatment	<i>Kras</i> Genotype	Tumours ( <i>n</i> )	<i>Kras</i> mutations
Urethane	WT	18	Q61R/L/H
	Het	26	Q61R/L/H
MNU	WT	5	G12D
	Het	21	G12D
None	LA2	12	G12D

Extended Data Table 2 | Mouse lung adenoma SNVs in established cancer driver genes and *Mtus1*

Chr	Position	Gene	Exon	Substitution	Consequence	Observed	Tumours	Validated	Validation Method*
12	112662237	<i>Akt1</i>	3	GGA>A	E40K	1	1024T10	Yes	Both
17	72603313	<i>Alk</i>	1	GGG>A	G133R	1	1024T8	Yes	Sequenom
18	34316299	<i>Apc</i>	16	TGT>A	Q2083X	1	1024T8	Yes	Inspection
4	133720797	<i>Arid1a</i>	3	GGA>A	S520F	1	1024T6	Yes	Sanger
4	133686649	<i>Arid1a</i>	15	GAC>G	D1287G	1	33T4	Yes	Sanger
17	5097671	<i>Arid1b</i>	3	GGG>A	P564L	1	1045T4	Yes	Sequenom
17	5337117	<i>Arid1b</i>	18	AGA>A	S1563F	1	1024T3	Yes	Sequenom
17	5337249	<i>Arid1b</i>	18	GGA>A	S1607F	1	1024T7	Yes	Sequenom
2	153393885	<i>Asxl1</i>	9	GGG>A	G296R	1	1024T2	Yes	Sequenom
2	153397578	<i>Asxl1</i>	11	GGC>A	P430S	1	1024T4	Yes	Sequenom
9	53460891	<i>Atm</i>	46	AGG>A	R2200K	1	75T3	Yes	Sequenom
9	53511883	<i>Atm</i>	13	AGA>A	E648K	1	1024T8	Yes	Sequenom
9	53518635	<i>Atm</i>	9	AGA>A	S367F	1	1024T4	Yes	Both
X	105875634	<i>Atrx</i>	9	GGG>A	G867R	1	1024T3	Yes	Sequenom
17	26190206	<i>Axin1</i>	8	AGC>A	A727T	1	1039T3	Yes	Inspection
11	101549022	<i>Brca1</i>	2	GGA>A	P25S	1	1024T7	Yes	Sequenom
5	150558455	<i>Brca2</i>	21	GGA>A	D2821N	1	1045T1	Yes	Sequenom
5	140882326	<i>Card11</i>	19	GGA>C	E856Q	1	1045T2	Yes	Inspection
9	44164145	<i>Cbl</i>	8	GGA>A	S401F	1	1024T9	Yes	Sanger
7	25286003	<i>Cic</i>	6	AGG>A	G291R	1	1012T3	Yes	Inspection
7	25287831	<i>Cic</i>	9	GGG>A	G481E	1	1024T5	Yes	Inspection
16	4085706	<i>Crebbp</i>	31	GGT>A	P1890S	1	1024T7	Yes	Sequenom
16	4094715	<i>Crebbp</i>	24	GGG>A	P1353S	1	1024T3	Yes	Sequenom
16	4117340	<i>Crebbp</i>	14	GGT>A	T933I	1	1024T1	Yes	Sequenom
17	33913569	<i>Daxx</i>	6	AGA>A	D596N	1	1024T2	Yes	Inspection
12	3899919	<i>Dnmt3a</i>	10	AGC>A	A352T	1	1045T4	Yes	Sequenom
15	81628398	<i>Ep300</i>	15	GGA>A	E974K	1	1024T5	Yes	Inspection
X	95428261	<i>Fam123b</i>	2	AGT>A	V84I	1	1024T4	Yes	Inspection
7	130196315	<i>Fgfr2</i>	9	CAG>G	C401R	1	1012T3	Yes	Inspection
5	33733951	<i>Fgfr3</i>	12	GGC>A	A539T	1	1024T1	Yes	Sequenom
5	33733706	<i>Fgfr3</i>	11	GGT>A	V482I	1	1024T5	Yes	Sanger
13	55160082	<i>Fgfr4</i>	7	GGC>A	A293V	1	1045T7	Yes	Inspection
5	147344556	<i>Flt3</i>	19	TGA>A	E789K	1	1045T4	Yes	Sequenom
6	88204692	<i>Gata2</i>	5	TAC>G	Y376C	1	1024T3	Yes	Sequenom
2	9874578	<i>Gata3</i>	3	GGA>A	E196K	1	1045T4	Yes	Sequenom
5	114952618	<i>Hnf1a</i>	7	GGG>A	P487S	1	1045T5	Yes	Sequenom
1	65161862	<i>Idh1</i>	7	GGC>A	G310D	1	1012T3	Yes	Inspection
19	29302040	<i>Jak2</i>	21	AGT>A	V1010I	1	1024T10	Yes	Sequenom
X	152268847	<i>Kdm5c</i>	19	CAG>T	Q902L	1	33T4	Yes	Sanger
X	152271108	<i>Kdm5c</i>	23	GAC>G	T1179A	1	35T1	Yes	Sanger
5	75647780	<i>Kit</i>	15	AGG>A	P728L	1	1024T2	Yes	Sequenom
4	55530863	<i>Klf4</i>	3	CAG>G	S83G	1	1800T2	Yes	Sequenom
13	111758076	<i>Map3k1</i>	11	AGA>A	D689N	1	1026T1	Yes	Inspection
12	81780619	<i>Map3k9</i>	1	GGG>A	G86S	1	1026T2	Yes	Sanger
12	81724480	<i>Map3k9</i>	10	GGT>A	T778I	1	1026T1	Yes	Sanger
12	81772793	<i>Map3k9</i>	2	AGG>A	R229K	1	75T1	Yes	Sanger
X	101294069	<i>Med12</i>	41	TAC>G	T1985A	1	33T2	Yes	Sequenom
19	6336766	<i>Men1</i>	3	AGG>A	G169R	1	1024T2	Yes	Sequenom
6	17562227	<i>Met</i>	19	CAA>C	K1196Q	1	1790T1	Yes	Sequenom
15	98852106	<i>Mll2</i>	32	GGG>A	P2569S	1	1024T4	Yes	Sequenom
15	98859560	<i>Mll2</i>	15	GGC>A	A1352T	1	1026T2	Yes	Sequenom
11	62343219	<i>Ncor1</i>	30	AGA>A	E1441K	1	1026T2	Yes	Sequenom
11	79425592	<i>Nf1</i>	13	TGT>A	C491Y	1	1024T3	Yes	Sequenom
3	98100211	<i>Notch2</i>	8	GGG>A	P426S	1	1045T2	Yes	Sequenom
4	44691909	<i>Pax5</i>	3	GGG>A	W112*	1	1024T10	Yes	Both
5	75181651	<i>Pdgfra</i>	15	AAT>T	N711I	1	309T1	Yes	Sequenom
5	75187929	<i>Pdgfra</i>	19	TGA>A	D877N	1	1045T1	Yes	Sequenom
17	20962623	<i>Ppp2r1a</i>	13	GGG>A	P523L	1	1024T7	Yes	Sequenom
13	63525046	<i>Ptch1</i>	17	AAC>G	N915S	1	1T2	Yes	Inspection
14	73206017	<i>Rb1</i>	22	GGA>A	S766F	1	1024T7	Yes	Inspection
14	73206083	<i>Rb1</i>	22	GGA>A	S744F	1	1024T5	Yes	Inspection
6	118164756	<i>Ret</i>	17	AGG>A	R970K	1	1024T5	Yes	Inspection
11	87731186	<i>Rnf43</i>	9	AGA>A	R371K	1	1024T8	Yes	Inspection
1	55012160	<i>Sf3b1</i>	6	GGT>A	T203I	1	1045T4	Yes	Sequenom
10	19011651	<i>Tnfrsf3</i>	2	GGT>A	T42I	1	1026T2	Yes	Sanger
8	41083460	<i>Mtus1</i>	2	GGG>A	W406*	2	1024T5, 1011T1	Yes	Both
8	41084181	<i>Mtus1</i>	2	CGT>A	T166M	1	1024T7	Yes	Both
8	41015397	<i>Mtus1</i>	7	GGG>A	G902R	1	1039T4	Yes	Both

\*Validation method: both = Sequenom MassArray and Sanger sequencing; inspection = manual inspection of alignments.

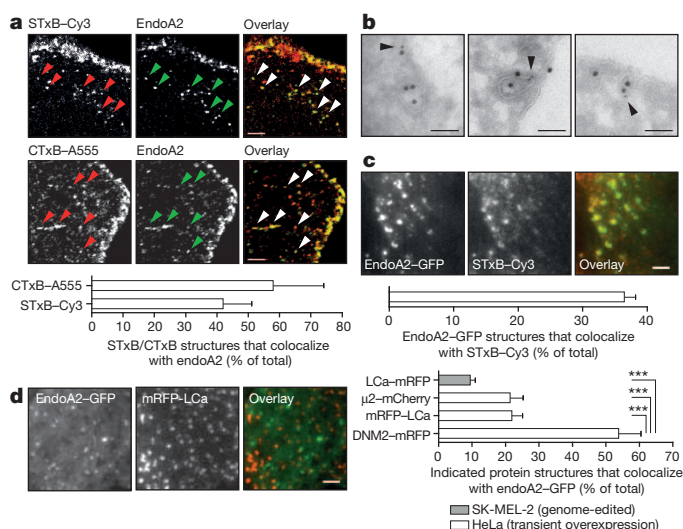
# Endophilin-A2 functions in membrane scission in clathrin-independent endocytosis

Henri-François Renard<sup>1,2,3\*</sup>, Mijo Simunovic<sup>4,5\*</sup>, Joël Lemièrre<sup>6,7\*</sup>, Emmanuel Boucrot<sup>8</sup>, Maria Daniela Garcia-Castillo<sup>1,2,3</sup>, Senthil Arumugam<sup>1,2,3</sup>, Valérie Chambon<sup>1,2,3</sup>, Christophe Lamaze<sup>2,3,9</sup>, Christian Wunder<sup>1,2,3</sup>, Anne K. Kenworthy<sup>10</sup>, Anne A. Schmidt<sup>11</sup>, Harvey T. McMahon<sup>12</sup>, Cécile Sykes<sup>6\*</sup>, Patricia Bassereau<sup>4\*</sup> & Ludger Johannes<sup>1,2,3</sup>

During endocytosis, energy is invested to narrow the necks of cargo-containing plasma membrane invaginations to radii at which the opposing segments spontaneously coalesce, thereby leading to the detachment by scission of endocytic uptake carriers<sup>1</sup>. In the clathrin pathway, dynamin uses mechanical energy from GTP hydrolysis to this effect<sup>2–4</sup>, assisted by the BIN/amphiphysin/Rvs (BAR) domain-containing protein endophilin<sup>5,6</sup>. Clathrin-independent endocytic events are often less reliant on dynamin<sup>7</sup>, and whether in these cases BAR domain proteins such as endophilin contribute to scission has remained unexplored. Here we show, in human and other mammalian cell lines, that endophilin-A2 (endoA2) specifically and functionally associates with very early uptake structures that are induced by the bacterial Shiga and cholera toxins, which are both clathrin-independent endocytic cargoes<sup>8</sup>. In controlled *in vitro* systems, endoA2 reshapes membranes before scission. Furthermore, we demonstrate that endoA2, dynamin and actin contribute in parallel to the scission of Shiga-toxin-induced tubules. Our results establish a novel function of endoA2 in clathrin-independent endocytosis. They document that distinct scission factors operate in an additive manner, and predict that specificity within a given uptake process arises from defined combinations of universal modules. Our findings highlight a previously unnoticed link between membrane scaffolding by endoA2 and pulling-force-driven dynamic scission.

Shiga toxin induces the clathrin-independent formation of endocytic plasma membrane invaginations as a first step in its entry into cells<sup>8</sup>. How this highly bent membrane domain is recognized by cellular machinery has remained unexplored. Here we screened an expression library of curvature-recognizing BAR domain proteins<sup>9</sup> for their localization to these structures. The screen was performed using the receptor-binding B-subunit of Shiga toxin (STxB) on cells with decreased levels of ATP to perturb active cellular machinery that is involved in the processing of these endocytic structures<sup>8,10</sup>. Out of 17 different BAR proteins, only Toca-1, Toca-3 and amphiphysin-2 scored positively (Extended Data Fig. 1a, yellow underlay). However, their short interfering RNA (siRNA)-mediated depletion did not affect STxB trafficking, hence these proteins were not studied further. In cells that expressed exogenous endoA2 (Extended Data Fig. 1a, green underlay), STxB-induced tubules were much shorter (Extended Data Fig. 1b), suggesting a functional crosstalk between endoA2 and the STxB uptake process. Because in non-ATP-depleted cells STxB trafficking to the Golgi was not detectably altered by expression of green fluorescent protein (GFP)-tagged endoA2 (Extended Data Fig. 1c), short tubules were probably the consequence of scission of longer tubules, rather than the inhibition of tubule formation.

At the plasma membrane and in very early uptake structures, STxB colocalized with endogenous endoA2 (Fig. 1a, b) or GFP-endoA2 (Fig. 1c, Extended Data Fig. 2a and Supplementary Video 1). Similar colocalization with endogenous endoA2 was observed for the GM1 glycosphingolipid-binding B-subunit of cholera toxin (CTxB) (Fig. 1a), which shares with STxB many aspects related to endocytic membrane bending<sup>11</sup>. Furthermore, the lifetime of structures containing endoA2 strongly increased in the presence of STxB (Extended Data Fig. 2b), and endoA2 was recruited



**Figure 1 | EndoA2 localization to endocytic pathways.** All conditions: incubation for 5 min at 37 °C (unless stated otherwise). **a**, BSC-1 cells with 50 nM STxB-Cy3 ( $n = 20$  cells) or 5 nM CTxB-Alexa555 (A555) ( $n = 50$  cells) for 3 min at 37 °C, and labelling for endoA2 (three independent experiments). **b**, Cryo-electron microscopy (representative of 50 images) on HeLa cells incubated with 0.5 μM STxB. Immunogold labelling: STxB, 15 nm gold particles; endoA2, 10 nm gold particles (arrowheads). **c**, HeLa cells transiently expressing endoA2-GFP incubated with 0.5 μM STxB-Cy3 and analysed by total internal reflection fluorescence microscopy (TIRFM) ( $n = 25$  cells, three independent experiments). **d**, Colocalization analysis by TIRFM of endoA2 with the indicated markers. Transient expression in HeLa cells: red fluorescent protein (RFP)-tagged dynamin-2 (DNM2-mRFP),  $n = 10$ ; mRFP-tagged clathrin light chain (mRFP-LCa),  $n = 8$ ; μ2-mCherry,  $n = 12$ . Genome-edited SK-MEL-2 cells: LCa-mRFP,  $n = 8$ . Two independent experiments. \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). Scale bars, 2 μm (**a**, **c**, **d**) and 100 nm (**b**). Data are mean  $\pm$  s.e.m.

<sup>1</sup>Institut Curie — Centre de Recherche, Endocytic Trafficking and Therapeutic Delivery group, 26 rue d'Ulm, 75248 Paris Cedex 05, France. <sup>2</sup>CNRS UMR3666, 75005 Paris, France. <sup>3</sup>U1143 INSERM, 75005 Paris, France. <sup>4</sup>Institut Curie — Centre de Recherche, Membrane and Cell Functions group, CNRS UMR 168, Physico-Chimie Curie, Université Pierre et Marie Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France. <sup>5</sup>The University of Chicago, Department of Chemistry, 5735 S Ellis Ave, Chicago, Illinois 60637, USA. <sup>6</sup>Institut Curie — Centre de Recherche, Biomimetic of Cell Movement group, CNRS UMR 168, Physico-Chimie Curie, Université Pierre et Marie Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France. <sup>7</sup>Université Paris Diderot, Sorbonne Paris Cité, 75205 Paris, France. <sup>8</sup>Institute of Structural and Molecular Biology, University College London & Birkbeck College, London WC1E 6BT, UK. <sup>9</sup>Institut Curie — Centre de Recherche, Membrane Dynamics and Mechanics of Intracellular Signaling group, 26 rue d'Ulm, 75248 Paris Cedex 05, France. <sup>10</sup>Vanderbilt School of Medicine, Department of Molecular Physiology and Biophysics, 718 Light Hall, Nashville, Tennessee 37232, USA. <sup>11</sup>CNRS, UMR7592, Institut Jacques Monod, Université Paris Diderot, Sorbonne Paris Cité, 15 rue Hélène Brion, 75205 Paris Cedex 13, France. <sup>12</sup>Medical Research Council, Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK.

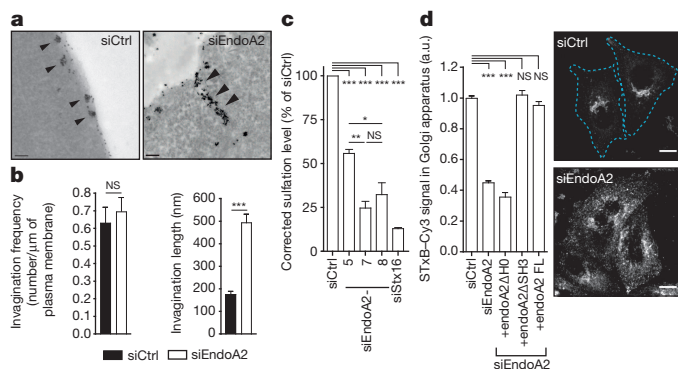
\*These authors contributed equally to this work.

to the plasma membrane by STxB (Extended Data Fig. 2c and Supplementary Video 2). These results demonstrated that endoA2 localized to sites of STxB and CTxB internalization.

Endophilins have classically been associated with the clathrin pathway (for example, see ref. 12), even if recent studies have indicated that this view may need to be broadened<sup>13,14</sup>. Notably, the overlap of endoA2 with fluorescent-protein-tagged clathrin or the  $\mu 2$  subunit of the clathrin adaptor AP2 was weak, both in transiently transfected HeLa cells, or genome-edited SK-MEL-2 cells (Fig. 1d and Extended Data Fig. 2d, e). STxB also only weakly colabelled with clathrin pathway markers in both cell systems (Extended Data Fig. 2g–j). By contrast, endoA2 extensively colocalized with dynamin (Fig. 1d and Extended Data Fig. 2f), as expected<sup>5,6</sup>.

In endoA2-depleted cells, short and apparently tubular STxB-containing invaginations were observed with sizes at the resolution limit of confocal light microscopy (Extended Data Fig. 3a). Using a nanogold conjugate of STxB that was specifically developed for this study (Extended Data Fig. 3b, c), it could be shown by electron microscopy that these structures were connected to the plasma membrane (Fig. 2a and Extended Data Fig. 3d). Their frequency did not increase after depletion of endoA2 (Fig. 2b, left), whereas their average length significantly increased (Fig. 2b, right). These findings were strongly in favour of a function for endoA2 in scission of STxB-induced endocytic plasma membrane invaginations.

Retrograde STxB trafficking to the *trans*-Golgi network (TGN) in endoA2-depleted cells was quantified using a STxB variant with tandem protein sulfation sites, termed STxB-Sulf<sub>2</sub>. With different siRNAs that had variable depletion efficiencies (Extended Data Fig. 4a), we found that sulfation of STxB-Sulf<sub>2</sub> in TGN membranes was inhibited in a dose-dependent manner with cellular endoA2 levels (Fig. 2c), while STxB cell-surface binding was not decreased (Extended Data Fig. 4b). Depletion of the TGN-localized tSNARE protein syntaxin-16 served as a benchmark treatment<sup>15</sup>. We concluded that endoA2 was required for efficient STxB trafficking into cells. A similar conclusion was reached when CTxB endocytosis was analysed in endoA2-depleted BSC-1 cells (Extended Data Fig. 4c).

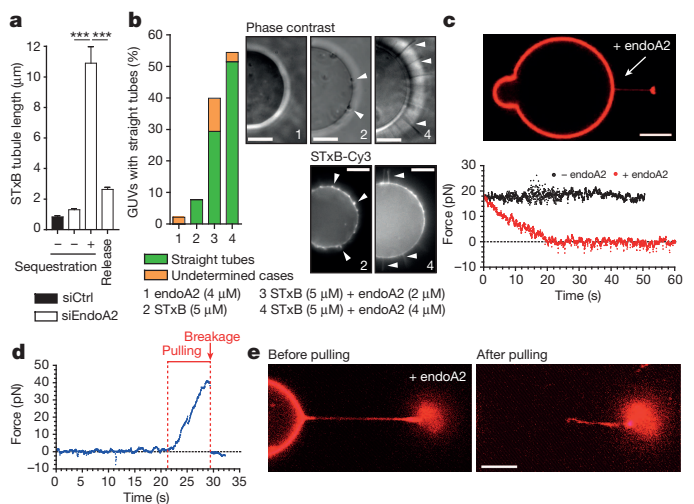


**Figure 2 | EndoA2 functions in Shiga toxin uptake.** **a**, Incubation for 5 min at 37 °C of control siRNA (siCtrl) or endoA2-depleted (siEndoA2) HeLa cells with nanogold-coupled STxB (arrowheads indicate invaginations; representative images of numbers of cells as shown in **b**). **b**, Frequency (siCtrl,  $n = 25$  cells; siEndoA2,  $n = 27$  cells) and length (siCtrl,  $n = 149$  tubules; siEndoA2,  $n = 138$  tubules) of STxB-containing invaginations on experiments as in **a**. NS, non significant. \*\*\* $P < 0.001$  (two-tailed Mann–Whitney  $U$  test). **c**, Sulfation analysis ( $n = 3$  independent experiments) on HeLa cells, using siRNAs targeting endoA2 (siEndoA2-5, 7 and 8), syntaxin-16 (siStx16) or control siRNA. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). **d**, Rescue experiment on endoA2-depleted HeLa cells. Incubation for 45 min at 37 °C with 50 nM STxB–Cy3. One arbitrary unit (a.u.) corresponds to  $78.7 \pm 1.2\%$  of STxB in Golgi area. Numbers of cells: siCtrl,  $n = 145$ ; siEndoA2,  $n = 190$ ; siEndoA2 plus endoA2 $\Delta$ H0,  $n = 46$ ; siEndoA2 plus endoA2 lacking the C-terminal SH3 domain (endoA2 $\Delta$ SH3),  $n = 31$ ; siEndoA2 plus full-length (FL) endoA2,  $n = 56$ ; two independent experiments. \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). Scale bars, 100 nm (**a**) and 10  $\mu\text{m}$  (**d**). Data are mean  $\pm$  s.e.m.

In endoA2-depleted cells, the inhibition of STxB transport to Golgi membranes could also be documented by immunofluorescence (Fig. 2d). Expression of siRNA-resistant endoA2 or a SH3 domain deletion mutant ( $\Delta$ SH3) rescued this phenotype (Fig. 2d and Extended Data Fig. 4d), suggesting that the function of endoA2 in STxB trafficking was independent of SH3 domain interaction with binding partners such as dynamin. By contrast, an amino-terminal amphipathic helix deletion mutant (endoA2 $\Delta$ H0) failed to rescue STxB trafficking (Fig. 2d and Extended Data Fig. 4d), probably owing to a role of this helix in membrane recruitment (see below). Correspondingly, endoA2 $\Delta$ SH3 expression reduced STxB-induced tubule length on ATP-depleted cells, similar to wild-type endoA2, whereas endoA2 $\Delta$ H0 failed to do so (Extended Data Fig. 4e), reinforcing the hypothesis of endoA2 participating in membrane scission.

Depletion of endoA2 did not affect transferrin internalization or recycling, the steady-state localization of the clathrin cargo proteins TGN46 and cation-independent mannose 6-phosphate receptor (CI-MPR)<sup>16,17</sup>, or anterograde transport of E-cadherin (Extended Data Fig. 5a–c). Clearly, it was the cellular entry of STxB and CTxB that was specifically altered under these conditions. By contrast, depletion of the  $\mu 2$  chain of AP2 had no effect on STxB trafficking, while transferrin uptake was efficiently blocked (Extended Data Fig. 5d–f).

To address the function of endoA2 in scission, we exploited the knocksideways system<sup>18</sup> to remove endoA2 acutely and reversibly from its normal site of action at the plasma membrane. HeLaM cells were used that expressed (1) the mitochondrial trap construct Mito-YFP-FRB (containing a yeast mitochondrial membrane protein, the yellow fluorescent protein (YFP) reporter and the FKBP–rapamycin-binding (FRB) domain), and (2) endoA2 tagged with both GFP and the rapamycin-binding protein FKBP. In cells with reduced ATP levels, the length of STxB-induced tubules was only slightly increased after depletion of endogenous endoA2 (Fig. 3a and Extended Data Fig. 6a), owing to the presence of the siRNA-resistant FKBP-tagged endoA2 fusion protein. The fusion protein was then sequestered on mitochondria by the addition of rapamycin, after which the length of STxB-containing tubules



**Figure 3 | Model membrane experiments.** **a**, Knocksideways technique on ATP-depleted HeLaM cells in control (siCtrl) or endoA2-depletion (siEndoA2) conditions. Incubations for 15 min at 37 °C with 0.5  $\mu\text{M}$  STxB. Mean  $\pm$  s.e.m. of the following numbers of tubules: siCtrl,  $n = 65$  (30 cells); without (–) sequestration,  $n = 101$  (52 cells); with (+) sequestration,  $n = 313$  (131 cells); release,  $n = 626$  (187 cells); two independent experiments. \*\*\* $P < 0.001$  (Dunn's multiple comparison test). **b**, Inverse emulsion technique in conditions 1–4, as indicated. Number of vesicles: 1,  $n = 46$ ; 2,  $n = 39$ ; 3,  $n = 17$ ; 4,  $n = 32$ ; three independent experiments. Arrows point to tubular structures. **c–e**, Nanotube tethers. **c**, Tube retraction force in presence (1  $\mu\text{M}$ ) or absence of endoA2. Representative of 21 experiments. **d**, **e**, Measurement of retraction force over time upon stepwise elongation at  $0.5 \mu\text{m s}^{-1}$  of an endoA2-scaffolded tube (**d**), and pulling-force-driven scission (**e**). Scale bars, 10  $\mu\text{m}$  (**b**) and 5  $\mu\text{m}$  (**c**, **e**).



significantly increased (Fig. 3a and Extended Data Fig. 6a). This effect was reversible, as rapamycin washout led to significantly shorter tubules (release condition). These results strongly reinforced the idea that endoA2 is involved in the scission of STxB-induced endocytic invaginations.

Controlled model membrane systems were then used to dissect aspects of endoA2 function with respect to membrane tubulation and scission. Cell-sized liposomes that reproduced a key aspect of natural membranes, that is, an asymmetric bilayer composition, were obtained via the inverse emulsion technique<sup>19</sup> (see Methods for leaflet compositions). To permit the acute introduction of endoA2 into the system, its topology was inverted with respect to the cellular situation: when indicated, STxB was included inside liposomes, whereas endoA2 was added to the external medium. Incubation with up to 4  $\mu\text{M}$  endoA2 did not lead to the formation of tubules (Fig. 3b, condition 1). By contrast, liposomes that contained STxB exhibited coiled tubular protrusions (condition 2). Addition of endoA2 to STxB-containing liposomes at concentrations of 2 or 4  $\mu\text{M}$  (Fig. 3b, conditions 3 or 4, respectively) caused tube straightening. Thus, endoA2 prompted the morphological transformation of STxB-induced tubular membranes.

For dynamic measurements, we turned to another setup in which optical tweezers were used to pull membrane tethers mechanically from giant unilamellar vesicles, held with a micropipette<sup>20</sup>. The force exerted by protein binding onto the tether membrane could be directly measured, tube radii controlled in a range from 10 to 100 nm, and tube elongation driven at controlled velocities. After endoA2 injection at concentrations as low as 1  $\mu\text{M}$  (protein concentration inside the injection pipette), the tube retraction force decreased to zero (Fig. 3c), indicating the formation of an external scaffold on the tether membrane, which probably explained the morphological transformation in the asymmetric bilayer system.

We did not observe spontaneous scission of endoA2-scaffolded tubes ( $n = 35$ ). However, when these were pulled further at elongation rates of  $\sim 0.5 \mu\text{m s}^{-1}$ , the tube retraction force increased (Fig. 3d), leading to breakage in 15 out of 18 cases (Fig. 3e). Elongation of bare tethers from fluid membranes, at rates as those used in this study, never leads to scission<sup>19,21</sup>, and we explicitly confirmed the resilience of bare tubes on forced elongation at an increased rate of  $\sim 10 \mu\text{m s}^{-1}$  ( $n = 2$ ). We speculate that interactions between endoA2 scaffold and the underlying membrane limit the flow of lipids during tube extension, thereby leading to increased membrane tension and spontaneous squeezing of the tube until the scission threshold radius has been reached.

The pulling speed that was required *in vitro* to induce the scission of endoA2-scaffolded tubes was very similar to the speed produced by microtubule-based molecular motors<sup>22</sup>, and microtubules have previously been implicated in the cellular entry of Shiga toxin<sup>23</sup>. Interference with microtubule-based functions before reduction of cellular ATP levels led to the loss of micrometre-long STxB-induced tubules (Extended Data Fig. 6b, c), while short ( $<200 \text{ nm}$ ) tubules persisted (Extended Data Fig. 7). These findings suggested that STxB-induced invaginations were recognized by molecular motors for the further processing into cells.

As indicated by a change in tube retraction force, the endoA2 $\Delta\text{H0}$  mutant formed a scaffold only at concentrations above 7  $\mu\text{M}$  (Extended Data Fig. 6d), seven times higher than the minimum concentration required for wild-type endoA2. Furthermore, the radius of the endoA2H0 scaffold was on average twice that of the wild-type endoA2 ( $23 \pm 10 \text{ nm}$ , mean  $\pm$  s.d.,  $n = 10$ , versus  $12 \pm 4 \text{ nm}$ ,  $n = 6$ , respectively). When pulled, endoA2 $\Delta\text{H0}$ -scaffolded tubes nevertheless underwent scission in 5 out of 5 cases (Extended Data Fig. 6e). The N-terminal amphipathic helix thus did not have a decisive role in scission by elongation of scaffolded membrane tubes. Rather, it seemed important for efficient recruitment of endoA2 to the membrane, as well as possibly affecting the molecular organization of the scaffold.

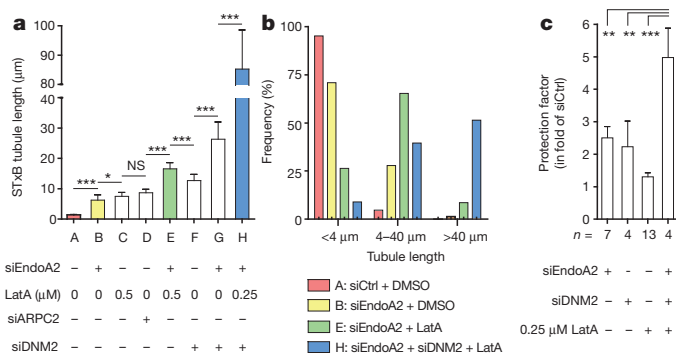
We next investigated the role of endoA2 together with other scission factors that have been shown before to play a part in STxB uptake. The first was actin<sup>10</sup>. EndoA2 and actin strongly and dynamically colocalized

(Extended Data Fig. 8a), and even in endoA2-depleted cells actin was localized on STxB-induced membrane invaginations (Extended Data Fig. 8b, c). This indicated that endoA2 was not required for actin recruitment to STxB uptake structures. Interestingly, the depletion of endoA2 or the depolymerization of actin with 0.5  $\mu\text{M}$  latrunculin-A led to a similar increase in the length of STxB-induced tubules, and the concomitant treatment had an additive effect (Fig. 4a and Extended Data Fig. 8d, e). Thus, both endoA2 and actin contributed independently to scission probability in STxB endocytosis. In cells that were depleted of ARPC2 (p34 subunit of the Arp2/3 complex), the length of STxB-induced tubules was similar to that observed in latrunculin-A-treated cells (Fig. 4a and Extended Data Fig. 8f), confirming the specificity of our observations.

After depletion of dynamin, STxB-induced tubules were longer than in control cells<sup>8</sup> (Fig. 4a), and endoA2 was clearly localized on these structures (Extended Data Fig. 9a, bottom). Tubules positive for endogenous endoA2 (Extended Data Fig. 9b) were seen only in the presence of STxB (compare  $\pm$  STxB panels), in line with our model membrane results. This finding documented that these tubules did not preexist but were indeed induced by STxB. Dynamin was still found in association with STxB uptake structures in endoA2-depleted cells (Extended Data Fig. 9c), indicating that in addition to endoA2, other proteins contribute to dynamin recruitment to membranes<sup>24</sup>. As for actin, concomitant depletion of dynamin and endoA2 had an additive effect on the length of STxB-induced tubules (Fig. 4a and Extended Data Fig. 9d).

Actin, dynamin and endoA2 thus independently contributed to the probability of scission in the STxB uptake pathway. Indeed, when dynamin and endoA2-depleted cells were concomitantly incubated with 0.25  $\mu\text{M}$  latrunculin-A, tubule length increased further, leading to an unprecedented tubulation phenotype throughout the cell population (Fig. 4a and Extended Data Fig. 9d). Importantly, when STxB-induced tubules were binned according to their length (Fig. 4b), a frequency shift was observed when going from unperturbed cells (condition A, red) to cells submitted to single (condition B, yellow), double (condition E, green) or triple (condition H, blue) perturbation. This data representation clearly illustrated that short tubules (most abundant in unperturbed cells) were progressively transformed into long tubules (most abundant in the triple-perturbation condition).

Protein biosynthesis inhibition is an established measure of Shiga toxin arrival in the cytosol<sup>25</sup>. Cells that were depleted of endoA2 or dynamin, or incubated with 0.25  $\mu\text{M}$  latrunculin-A, were only weakly protected against Shiga-like toxin-1 (STx-1; note that STx-1 and Shiga toxin differ by only one residue) (Fig. 4c and Extended Data Fig. 10a–c). These



**Figure 4 | Additive effects of scission factors.** **a**, Incubation of HeLa cells for 5 min at 37 °C with 0.5  $\mu\text{M}$  STxB-Cy3. Determination of tube length on fixed cells (number of tubules per condition: A,  $n = 254$ ; B,  $n = 79$ ; C,  $n = 94$ ; D,  $n = 146$ ; E,  $n = 118$ ; F,  $n = 338$ ; G,  $n = 131$ ; H,  $n = 115$ ); two independent experiments. NS, non significant, \* $P < 0.05$ , \*\*\* $P < 0.001$  (Dunn's multiple comparison test). LatA, latrunculin-A. **b**, Frequency distribution of tubules according to length (conditions as in **a**). **c**, Intoxication analysis. Protection factors on 4–13 independent experiments per condition. \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). Data are mean  $\pm$  s.e.m.

individual treatments thus failed to impose a transport block that would shut down access to the cytosol to levels that were required to be robustly detected by this extremely sensitive method. By contrast, triple treatment induced a substantial shift to the right of the intoxication curves (Extended Data Fig. 10d), such that 5.0-fold ( $\pm 0.9$ ,  $n = 4$ ) more STx-1 was required to achieve the same level of protein biosynthesis inhibition as in control conditions (see quantification in Fig. 4c). Clearly, the scission factors endoA2, actin and dynamin needed to be inhibited simultaneously for a substantial protection of cells against STx-1.

Our study reveals a novel function of endoA2 in clathrin-independent endocytosis of Shiga and cholera toxin, and suggests more generally that this BAR domain protein operates preferentially on non-clathrin sites of uptake (see ref. 26). Our data also provide strong evidence for a function of endoA2 in the scission process. In the context of the Shiga toxin model, endoA2 is, to our knowledge, the first cellular factor that was identified to recognize specifically the mechanical signal transmitted by the toxin to the cytosol: a highly curved plasma membrane domain. Unexpectedly, our findings suggest that actin<sup>10</sup> and dynamin<sup>8</sup> contribute in an additive manner with endoA2 to scission in Shiga toxin uptake. They probably act on the same endocytic invaginations and not on parallel uptake events, since the overall length of Shiga-toxin-induced plasma membrane tubules strongly increased in cells in which all three scission factors were inhibited. Such additive function in scission is unexpected and suggests that specificity within given endocytic processes is the result of defined cocktails of modular machinery. This hypothesis does not exclude a functional interaction between scission factors<sup>27,28</sup>. An exciting new modality emerges from our observations that scaffolding by endoA2 sensitizes tether membranes for pulling-force-induced scission. This finding should stimulate further research into the function of cytoskeleton-based molecular motors in scission and the mechanisms by which scaffolding by BAR domain proteins prime membranes to this effect.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 15 August; accepted 14 November 2014.**

**Published online 17 December 2014.**

1. Campelo, F. & Malhotra, V. Membrane fission: the biogenesis of transport carriers. *Annu. Rev. Biochem.* **81**, 407–427 (2012).
2. Shnyrova, A. V. *et al.* Geometric catalysis of membrane fission driven by flexible dynamin rings. *Science* **339**, 1433–1436 (2013).
3. Morlot, S. *et al.* Membrane shape at the edge of the dynamin helix sets location and duration of the fission reaction. *Cell* **151**, 619–629 (2012).
4. Faelber, K. *et al.* Crystal structure of nucleotide-free dynamin. *Nature* **477**, 556–560 (2011).
5. Sundborger, A. *et al.* An endophilin-dynamin complex promotes budding of clathrin-coated vesicles during synaptic vesicle recycling. *J. Cell Sci.* **124**, 133–143 (2011).
6. Neumann, S. & Schmid, S. L. Dual role of BAR domain-containing proteins in regulating vesicle release catalyzed by the GTPase, dynamin-2. *J. Biol. Chem.* **288**, 25119–25128 (2013).
7. Howes, M. T., Mayor, S. & Parton, R. G. Molecules, mechanisms, and cellular roles of clathrin-independent endocytosis. *Curr. Opin. Cell Biol.* **22**, 519–527 (2010).
8. Römer, W. *et al.* Shiga toxin induces tubular membrane invaginations for its uptake into cells. *Nature* **450**, 670–675 (2007).
9. Mim, C. & Unger, V. M. Membrane curvature and its generation by BAR proteins. *Trends Biochem. Sci.* **37**, 526–533 (2012).
10. Römer, W. *et al.* Actin dynamics drive membrane reorganization and scission in clathrin-independent endocytosis. *Cell* **140**, 540–553 (2010).
11. Ewers, H. *et al.* GM1 structure determines SV40-induced membrane invagination and infection. *Nature Cell Biol.* **12**, 11–18 (2010).
12. Milosevic, I. *et al.* Recruitment of endophilin to clathrin-coated pit necks is required for efficient vesicle uncoating after fission. *Neuron* **72**, 587–601 (2011).

13. Llobet, A. *et al.* Endophilin drives the fast mode of vesicle retrieval in a ribbon synapse. *J. Neurosci.* **31**, 8512–8519 (2011).
14. Kononenko, N. L. *et al.* Clathrin/AP-2 mediate synaptic vesicle reformation from endosome-like vacuoles but are not essential for membrane retrieval at central synapses. *Neuron* **82**, 981–988 (2014).
15. Amessou, M. *et al.* Syntaxin 16 and syntaxin 5 control retrograde transport of several exogenous and endogenous cargo proteins. *J. Cell Sci.* **120**, 1457–1468 (2007).
16. Draper, R. K., Goda, Y., Brodsky, F. M. & Pfeffer, S. R. Antibodies to clathrin inhibit endocytosis but not recycling to the *trans* Golgi network *in vitro*. *Science* **248**, 1539–1541 (1990).
17. Nesterov, A., Carter, R. E., Sorkina, T., Gill, G. N. & Sorkin, A. Inhibition of the receptor-binding function of clathrin adaptor protein AP-2 by dominant-negative mutant  $\mu 2$  subunit and its effects on endocytosis. *EMBO J.* **18**, 2489–2499 (1999).
18. Robinson, M. S., Sahlender, D. A. & Foster, S. D. Rapid inactivation of proteins by rapamycin-induced rerouting to mitochondria. *Dev. Cell* **18**, 324–331 (2010).
19. Campillo, C. *et al.* Unexpected membrane dynamics unveiled by membrane nanotube extrusion. *Biophys. J.* **104**, 1248–1256 (2013).
20. Cuvelier, D., Derenyi, I., Bassereau, P. & Nassoy, P. Coalescence of membrane tethers: experiments, theory, and applications. *Biophys. J.* **88**, 2714–2726 (2005).
21. Evans, E. & Yeung, A. Hidden dynamics in rapid changes of bilayer shape. *Chem. Phys. Lipids* **73**, 39–56 (1994).
22. Toba, S., Watanabe, T. M., Yamaguchi-Okimoto, L., Toyoshima, Y. Y. & Higuchi, H. Overlapping hand-over-hand mechanism of single molecular motility of cytoplasmic dynein. *Proc. Natl Acad. Sci. USA* **103**, 5741–5745 (2006).
23. Hehnl, H., Sheff, D. & Stammes, M. Shiga toxin facilitates its retrograde transport by modifying microtubule dynamics. *Mol. Biol. Cell* **17**, 4379–4389 (2006).
24. Meinecke, M. *et al.* Cooperative recruitment of dynamin and BIN/Amphiphysin/Rvs (BAR) domain-containing proteins leads to GTP-dependent membrane scission. *J. Biol. Chem.* **288**, 6651–6661 (2013).
25. Johannes, L. & Römer, W. Shiga toxins - from cell biology to biomedical applications. *Nature Rev. Microbiol.* **8**, 105–116 (2010).
26. Boucrot, E. *et al.* Endophilin marks and controls a clathrin-independent endocytic pathway. *Nature* <http://dx.doi.org/10.1038/nature14067> (this issue).
27. Itoh, T. *et al.* Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. *Dev. Cell* **9**, 791–804 (2005).
28. Roux, A., Uyhazi, K., Frost, A. & De Camilli, P. GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission. *Nature* **441**, 528–531 (2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We would like to acknowledge the following people for help in experiments and providing materials or expertise: A. Berthier, L. Cabanié, K. Carvalho, P. de Camilli, C. Day, D. Drubin, A. El Marjou, V. Fraissier, A. Gautreau, T. Kirchhausen, L. Leconte, C. Merrifield, G. Montagnac, P. Paul-Gilloteaux, M. S. Robinson, L. Sengmanivong and E. Smythe. The facilities as well as scientific and technical assistance from staff in the PICT-IBISA/Nikon Imaging Centre at Institut Curie-CNRS and the France-Biolmaging infrastructure (ANR-10-INSB-04) are acknowledged. This work was supported by grants from the Agence Nationale pour la Recherche (ANR-09-BLAN-283 to L.J. and C.S., ANR-10-LBX-0038 to C.L., ANR-11-BSV2-014-03 to L.J. and P.B., ANR-12-BSV5-0014 to C.S.), the Indo-French Centre for the Promotion of Advanced Science (project no. 3803, L.J.), Marie Curie Actions — Networks for Initial Training (FP7-PEOPLE-2010-ITN, L.J.), European Research Council advanced grant (project 340485, L.J.), Marie Curie International Reintegration Grant (FP7-RG-277078, C.W.), the Royal Society (RG120481, E.B.), Fondation ARC pour la Recherche sur le Cancer (DEQ20120323737, C.S.), National Institutes of Health (R01 GM106720, A.K.K.), La Ligue contre le Cancer, Comité de Paris (RS08/75-89, A.A.S.), and by fellowships from Fondation ARC pour la Recherche sur le Cancer (H.-F.R., J.L. and M.-D.G.-C.), AXA Research Funds (J.L. and M.-D.G.-C.), the Biological Sciences Research Council (David Phillips Research Fellowship to E.B.), Chateaubriand fellowship and the France and Chicago Collaborating in the Sciences grant (M.S.). The Johannes, Lamaze, Sykes and Bassereau teams are members of Labex CellTisPhyBio (11-LBX-0038) and of Idex Paris Sciences et Lettres (ANR-10-IDEX-0001-02 PSL). P.B.'s group belongs to the French research consortium CellTis.

**Author Contributions** The following datasets were contributed by the indicated authors: Figs 1b–d, 2–4 and Extended Data Figs 1–3, 4a, b, d, e, 5–10 (H.-F.R.); Fig. 3c–e and Extended Data Fig. 6d, e (M.S.); Fig. 3b (J.L.); Fig. 1a and Extended Data Fig. 4c (E.B.); Fig. 4c and Extended Data Fig. 10 (M.-D.G.-C.); Fig. 2a and Extended Data Fig. 3d (S.A. and V.C.); Fig. 1b and Extended Data Figs 3b, c and 7 (V.C.). C.L., C.W., A.K.K., A.A.S. and H.T.M. provided technical support and conceptual advice. C.S., P.B. and L.J. provided direction and guidance. L.J. conceived the initial design of the study and wrote the first draft. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.J. ([ludger.johannes@curie.fr](mailto:ludger.johannes@curie.fr)).

## METHODS

**Antibodies and other reagents.** The following antibodies were purchased from the indicated suppliers: rabbit polyclonal anti-endoA2 for western blotting (Bethyl, A302-349A, 1:2,000) or immunofluorescence and cryo-electron microscopy (Santa Cruz Biotechnology, sc-25495, 1:200 for immunofluorescence and 1:10 for cryo-electron microscopy); rabbit polyclonal anti-STxB (Covalab, 1:100 for cryo-electron microscopy); mouse monoclonal anti-clathrin heavy chain for western blotting (BD Biosciences, 610500, 1:5,000); mouse monoclonal anti- $\alpha$ -tubulin (Sigma, T5168, 1:5,000 for western blotting); rabbit polyclonal anti-syntaxin-16 (Synaptic Systems, 110 163, 1:1,000 for western blotting); rabbit monoclonal anti-giantin (Institut Curie, recombinant proteins platform, A-R-R#05, 1:100 for immunofluorescence); sheep polyclonal anti-TGN46 (Serotec, AHP500G, 1:200 for immunofluorescence); mouse monoclonal anti-CI-MPR (Abcam, ab2733, 1:200 for immunofluorescence); rabbit polyclonal anti-mCherry (Institut Curie, Recombinant proteins platform, A-P-R#13, 1:200 for immunofluorescence); mouse monoclonal anti-GFP (Roche Applied Sciences, 11 814 460 001, 1:1,000 for western blotting); mouse monoclonal anti-dynamin Hudy 1 (Upstate, 05-319, 1:200 for immunofluorescence); mouse monoclonal anti-dynamin (BD Biosciences, 610246, 1:1,000 for western blotting); rabbit polyclonal anti-ARPC2 (Millipore, 07-227, 1:1,000 for western blotting); rabbit polyclonal anti-DYNC1H1 (Proteintech, 12345-1-AP, 1:200 for western blotting); secondary antibodies conjugated to Alexa488, Cy3, Cy5, AMCA or horseradish peroxidase (HRP) (Beckman Coulter and Invitrogen). The mouse monoclonal anti-STxB antibody 13C4 was purified from hybridoma cells (ATCC CRL-1794), and the mouse monoclonal anti-clathrin heavy chain antibody X22 (used at 1:1,000 for immunofluorescence) was a gift from E. Smythe. 2-deoxy-D-glucose, sodium azide, latrunculin-A and nocodazole were purchased from Sigma.

**Cell culture.** HeLa and BSC-1 cells were grown at 37 °C under 5% CO<sub>2</sub> in DMEM high glucose glutamax (Invitrogen) supplemented with 10% FCS, 0.01% penicillin-streptomycin and 5 mM pyruvate. Genome-edited SK-MEL-2 cells expressing mRFP-LCa were provided by D. Drubin, and grown at 37 °C under 5% CO<sub>2</sub> in DMEM/F12 (Invitrogen) supplemented with 10% FCS, 0.01% penicillin-streptomycin and 5 mM pyruvate. HeLaM cells stably expressing Mito-YFP-FRB were provided by M. S. Robinson. These were grown at 37 °C under 5% CO<sub>2</sub> in DMEM high glucose glutamax (Invitrogen) supplemented with 10% FCS, 0.01% penicillin-streptomycin, 5 mM pyruvate and 137.5  $\mu$ g ml<sup>-1</sup> hygromycin B. HeLaM Mito-YFP-FRB cells stably expressing C-terminally GFP-FKBP-tagged rat endoA2 were generated for this study (see below) and grown in the same medium as the mother cells, supplemented with 0.5 mg ml<sup>-1</sup> G418.

**Depletion of cellular ATP.** Cellular ATP was depleted as previously described<sup>8,29</sup>. In brief, cells were incubated for 15–20 min at 37 °C in PBS<sup>++</sup> supplemented with 10 mM 2-deoxy-D-glucose and 10 mM NaN<sub>3</sub>.

**DNA constructs and transfection.** Expression plasmids for C-terminally GFP-tagged rat endoA2 (P. De Camilli), mCherry-tagged  $\mu$ 2 subunit of AP2 (C. Merrifield), mRFP-tagged dynamin-2 or clathrin light chain (LCa) (T. Kirchhausen) were provided by the indicated colleagues.

A bicistronic vector encoding SBP-mCherry-E-cadherin and KDEL-streptavidin, used for the anterograde E-cadherin transport assay, was provided by F. Perez, and a plasmid encoding LifeAct-mCherry<sup>30,31</sup> by P. Chavrier.

GFP-tagged mutants of endoA2 ( $\Delta$ H0 lacking N-terminal amphipathic H0 helix, and  $\Delta$ SH3 lacking C-terminal SH3 domain) were obtained by PCR from full-length endoA2.

For knocksideways with endoA2, a C-terminally GFP-FKBP-tagged rat endoA2 construct was engineered. For this, rat endoA2-GFP and the FKBP sequence bearing a stop codon were amplified by PCR from plasmids (for amplification of endoA2-GFP fragment, forward primer was 5'-ATACTTAAGATGTCGGTGGCGGGGCGTGAAG-3' and reverse primer was 5'-TTCCACCTGCACTCCCATCCCTCCG CCCTGTACAGCTCGTCCAT-3'; for amplification of FKBP fragment, forward primer was 5'-ATGGACGAGCTGTACAAGGGCGGAGGATGGGAGTGCA GGTGGAA-3' and reverse primer was 5'-AGTGCGGCCGCTTATCCAGTTT TAGAAGCTC-3'). The PCR fragments were designed in such a way that they possessed overlapping sequences. This allowed obtaining an endoA2-GFP-FKBP fragment via a third PCR on the two previous fragments with appropriate primers (forward primer from endoA2-GFP amplification: 5'-ATACTTAAGATGTCGGTGGCGGGGCTGAAG-3'; reverse primer from FKBP fragment amplification: 5'-AGTGCGGCCGCTTATCCAGTTT TAGAAGCTC-3'). After insertion in a pIRESneo2 vector between AflIII and NotI restriction sites, clones were validated by sequencing.

For immunofluorescence and live-cell imaging experiments, plasmids were transfected with FuGene 6 (Promega) according to the manufacturer's instructions, or using the classical calcium phosphate procedure<sup>32</sup>. Cells were used for experiments 16–24 h after transfection. For the production of a HeLaM Mito-YFP-FRB cell line stably expressing endoA2-GFP-FKBP, the pIRESneo2/endoA2-GFP-FKBP plasmid was transfected by electroporation, and clones were selected with 0.5 mg ml<sup>-1</sup> G418.

**RNA interference.** Most siRNAs used in this study were purchased from Qiagen and transfected with HiPerFect (Qiagen) according to manufacturer's instructions. Experiments were performed 72 h after siRNA transfection, where protein depletion efficiency was maximal as shown by immunoblotting analysis with specific antibodies (routinely 80–90%). For most experiments, cells were replated 24 h before use, according to the need of the experiment. AllStars Negative Control siRNA served as a reference point. The depletion of endoA2 was achieved with three different sequences at final concentrations between 2 and 10 nM: 5 (SI03057250: 5'-CACCAGCAAGGCGGTGACAGA-3'), 7 (SI03073931: 5'-CATGCTCAACA CGGTGTCCAA-3') and 8 (SI03108049: 5'-TACACTAGCGCTGACTCCCAA-3'). For syntaxin-16 depletion, a custom-made siRNA sequence was used<sup>15</sup> at concentrations between 2 and 10 nM (5'-AAGCAGCGATTGGTGTGACAA-3'); for dynamin-2, a single siRNA at 10 nM (SI02654687: 5'-CTGACGCTCATCTTCTC AAAA-3'); for  $\alpha$ -adaptin, a single siRNA sequence<sup>33</sup> at concentrations between 10 and 40 nM (5'-AAATGGCGGTGGTGTGCGGCTC-3', Dharmacon); for ARPC2, a smartpool of 4 siRNA sequences (L-012081-00, Dharmacon: 5'-CCATGTATG TTGAGTCTAA-3', 5'-GCTCTAAGGCCTATATTCA-3', 5'-GGACAGAGTCA CAGTAGTC-3' and 5'-GTACGGGAGTTTCTTGGA-3') at 40 nM.

For endoA2 depletion from BSC-1 cells, a single all Stealth siRNA was used (HSS109705, Invitrogen: 5'-GCTGACCAACCAGATCGATGAGAAC-3'). Cells seeded on 12 mm coverslips placed in 24-well plates (for microscopy) or on glass-bottom 96-well plates (plate-reader experiments) were transfected twice (on days 1 and 2) with oligofectamine (Life Technologies) complexed to 8 or 1.6 pmol, respectively, of HSS109705 siRNA and analysed on day 4 (72 h after the first transfection).

**Recombinant proteins.** Recombinant wild-type STxB, STxB-Cys and STxB-Sulf<sub>2</sub> were purified from bacterial periplasmic extracts as previously described<sup>34</sup>. C-terminally Strep-tagged mouse endoA2 and endoA2 $\Delta$ H0 (N-terminal amphipathic helix deletion mutant) were expressed in bacteria and purified on a Strep-Tactin column (IBA), as previously described<sup>35</sup>. See Supplementary Information for more details.

**Light microscopy.** For immunofluorescence studies, cells were maintained at 37 °C during the full duration of the experiment and during fixation (4% paraformaldehyde) to preserve the integrity of STxB-induced tubules. After quenching with 50 mM NH<sub>4</sub>Cl and permeabilization with saponin (0.5% saponin, 2% BSA in PBS), cells were incubated with primary and secondary antibodies, and mounted with Mowiol.

Fixed samples were imaged with a Nikon A1R confocal microscope equipped with a CFI Plan Apo VC 60 $\times$  numerical aperture (NA) 1.4 oil immersion objective, when not specified otherwise. Wide field images were acquired on a Leica DM 6000B epifluorescence inverted microscope equipped with a HCX PL Apo 63 $\times$  NA 1.40 oil immersion objective and an EMCCD camera (Photometrics CoolSNAP HQ). Super resolution images were acquired on a N-SIM (Structured Illumination Microscopy, Nikon) equipped with a CFI Apo TIRF 100 $\times$  NA 1.49 oil immersion objective and an EMCCD camera (Photometrics CoolSNAP HQ2).

For live-cell imaging, cells were grown to subconfluence on FluoroDish chambers with integrated glass coverslips (World Precision Instruments). All observations were made at 37 °C and 5% CO<sub>2</sub>. Different imaging devices were used. Plasma membrane images were acquired on a TIRF video microscope (Nikon) equipped with a CFI Apo TIRF 100 $\times$  NA 1.49 oil objective and an EMCCD camera (Photometrics HQ2). Other live-cell images were acquired on spinning disk confocal devices (Nikon) equipped with EMCCD cameras (Photometrics CoolSNAP HQ2). Montages, kymographs and videos were prepared with ImageJ or Fiji (NIH) and MetaMorph Software.

**BAR domain screening.** After 24 h transfection of HeLa cells with fluorescent-protein-tagged constructs, cells were ATP-depleted for 20 min, incubated for 20 min with 1  $\mu$ M STxB-Cy3 or STxB-Alexa488, fixed at 37 °C, mounted, and viewed by confocal microscopy. The initial phenotype that was scored was BAR domain protein localization on STxB-induced plasma membrane invaginations.

**Electron microscopy.** Two different cell embedding techniques were used: cryo-protection and epon. For further details, see Supplementary Information.

**STxB sulfation.** The sulfation assay was performed as previously described<sup>36</sup>. For further details, see Supplementary Information.

**CTxB internalization.** Alexa555-labelled CTxB (Life Technologies) was used at 0.9–5 nM (50–285 ng ml<sup>-1</sup>) and incubated on cells directly at 37 °C and 5% CO<sub>2</sub> (without pre-incubation on ice) for up to 10 min. After incubation, samples destined for microscopy were quickly washed once and fixed at 37 °C with pre-warmed 3.7% paraformaldehyde (PFA) in PBS (20 min, 37 °C). Samples were then washed (twice with PBS and once with 50 mM NH<sub>4</sub>Cl in PBS), immunostained and imaged on a laser-scanning confocal microscope (TCS Sp5 AOBs; Leica). For CTxB uptake assay, cells grown in 96-well glass bottom black plates (Thermo Scientific) were incubated with 5 nM toxin for 10 min at 37 °C, moved on ice to stop further endocytosis, washed once with ice-cold PBS to remove unbound toxin and twice with an ice-cold low pH solution (300 mM NaCl, 5 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 0.2 M acetic acid, pH 2.5, as described previously<sup>37,38</sup>), which removed 90  $\pm$  6% of



the cell surface signals ( $n = 3$ ,  $P < 0.0001$ , two tailed  $t$ -test) as compared to samples incubated with the toxin on ice (surface staining). The samples were then further washed twice with ice-cold PBS (to increase the pH back to 7.4) and finally fixed on ice with 3.7% PFA for 30 min. Cells were then washed twice with PBS and once with 50 mM NH<sub>4</sub>Cl in PBS and incubated with 4',6-diamidino-2-phenylindole (DAPI) to label their DNA (used to normalize the number of cells in between samples and experiments). Signals from internalized CTxB-Alexa555 (protected from the acid wash) and DAPI were measured using a plate reader (FLUOstar Optima, BMG).

**Intoxication.** The intoxication assay with STx-1 holotoxin was performed as previously described<sup>39</sup>. For further details, see Supplementary Information.

**Knocksideways.** The knocksideways technique was adapted to endoA2, as described<sup>40</sup>. For further details, see Supplementary Information.

**Flow cytometry.** siRNA transfected or PMP-2-treated cells (as control) were detached using 4 mM EDTA and incubated for 30 min on ice with transferrin-Alexa647 (TF-A647; 10  $\mu$ g ml<sup>-1</sup>) and STxB-cysteine-Alexa488 (1  $\mu$ M) in PBS<sup>++</sup> containing 0.2% BSA. After washing in ice-cold PBS, fluorescence was measured with a LSR-II flow cytometer (BD Biosciences). Single-stained samples were used to verify that the fluorescence from each fluorophore was only detected in the expected channel.

**Transferrin endocytosis.** The endocytosis assay was performed essentially as previously described<sup>36</sup>. For further details, see Supplementary Information.

**Transferrin recycling.** Cells were incubated for 40 min at 37 °C with biotinylated transferrin with a disulfide bond (TF-SS-biot; 60  $\mu$ g ml<sup>-1</sup>) in PBS<sup>++</sup> supplemented with 5 mM glucose and 0.2% BSA. After washing, cells were aliquoted (150,000 cells per data point) and placed for the indicated times at 37 °C in PBS<sup>++</sup> supplemented with 5 mM glucose and 0.2% BSA in the presence of a 50-fold molar excess of non-biotinylated holo-Tf (Sigma-Aldrich). The cells were placed on ice, washed in PBS<sup>++</sup>, and lysed in 10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 1 mM EDTA, 0.2% BSA, 0.1% SDS, 1% Triton X-100. The amount of cell-associated biotinylated transferrin was determined by ELISA, as in the endocytosis assay.

**Anterograde E-cadherin transport.** The retention using selective hooks (RUSH) system was used to quantify anterograde transport of E-cadherin, as previously described<sup>41</sup>. For further details, see Supplementary Information.

**Asymmetric liposomes produced by inverted emulsion.** Lipids and reagents. Egg phosphatidylcholine (EPC), cholesterol, L- $\alpha$ -phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) and brain total lipid extract were purchased from Avanti Polar Lipids. Ceramide trihexosides (Gb3) was purchased from Matreya LLC. All chemicals, including mineral oil were purchased from Sigma Aldrich unless specified otherwise.

Buffer composition. Inside buffer: 25 mM HEPES, pH 7.4, 100 mM NaCl, 0.5 mM EGTA and 75 mM sucrose. Outside buffer: 25 mM HEPES, pH 7.4, 100 mM NaCl, 0.5 mM EGTA, 0.05 mg ml<sup>-1</sup>  $\beta$ -casein and 75 mM glucose.

Liposome preparation. Asymmetric liposomes were obtained using an inverted emulsion technique<sup>42,43</sup>. Lipids were dissolved in mineral oil at a total concentration of 0.5 mg ml<sup>-1</sup>, with a composition of egg phosphatidylcholine:cholesterol:Gb3 at a molar ratio of 65:30:5 (inner leaflet) or of brain total lipid extract: PtdIns(4,5)P<sub>2</sub> at a mass ratio of 95:5 (outer leaflet). A volume of 2.5 ml of each oil-lipid mixture was sonicated in a bath at 35 °C for 30 min at a power of 30 W, cooled and stored for 1 day at room temperature. A volume of 30  $\mu$ l of the outer leaflet oil-lipid mixture was poured in a tube on top of the same volume of the outside buffer. The tube was left at room temperature for at least 2 h. In the meantime, an emulsion was obtained by mixing a volume of 250  $\mu$ l of inner leaflet oil-lipid mixture with 1.25  $\mu$ l of inside buffer (with or without proteins) and by pumping back-and-forth with a syringe. Then, 50  $\mu$ l of this emulsion was slowly added to the top phase of the tube and immediately centrifuged at 250g for 5 min. Asymmetric liposomes were thus obtained at the bottom of the tube and were collected by removing the upper oil phase.

Liposome observation. Liposomes were observed with an IX70 Olympus inverted microscope equipped with a Olympus 100 oil-immersion phase-contrast objective, numerical aperture (NA) 1.35 (Olympus). Fluorescent Shiga toxin was excited by a 200 W mercury lamp (OSRAM). Images were acquired with a cooled CCD camera CoolSNAP ES (Photometrics).

**Preparation of giant unilamellar vesicles.** Brain total lipid extract, 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-*N*-[biotinyl(polyethyleneglycol)-2000] (DSPE-PEG(2000)-biotin), and PtdIns(4,5)P<sub>2</sub> were purchased from Avanti Polar Lipids. BODIPY-TR-C5-ceramide was obtained from Molecular Probes. All reagents used to make buffers and  $\beta$ -casein from bovine milk (>99%) were purchased from Sigma-Aldrich.

Giant unilamellar vesicles (GUVs) were prepared by electroformation on platinum wires. Approximately 5  $\mu$ l of the lipid mix, containing ~95% brain total extract, 5% PtdIns(4,5)P<sub>2</sub>, ~0.3% DSPE-PEG(2000)-biotin and 0.5% BODIPY-TR-C5-ceramide (all molar fractions), was deposited on platinum wires at 3 g l<sup>-1</sup>, and dried under vacuum for 1 h. GUVs were grown overnight in a mixture of 200 mM sucrose, 50 mM NaCl, 20 mM HEPES, pH 7.4, at  $v = 500$  Hz,  $I = 0.25$  V,  $t = 4$  °C.

**Tether-pulling experiments and data analysis.** An experimental chamber, consisting of two parallel coverslips, spaced a few millimetres from each another, was immersed in a 5 g l<sup>-1</sup> solution of  $\beta$ -casein for 30 min to passivate the glass surface. The chamber was filled with 5–10  $\mu$ l of the electroformation solution and 200  $\mu$ l buffer solution (200 mM glucose, 50 mM NaCl, 20 mM HEPES, pH 7.4). The chamber was sealed with oil after ~10 min to prevent evaporation. GUVs with visible fluctuations were aspirated in a micropipette. Membrane tension was controlled via the aspiration pressure. A tether was created by bringing in contact the GUV with a streptavidin-coated polystyrene bead, approximately 3  $\mu$ m in diameter (SpheroTec), trapped with optical tweezers. Using another micropipette, endoA2 was injected in the vicinity of the GUV (at bulk concentrations in the micropipette 0.3–5  $\mu$ M per monomer for the wild type and 5–7  $\mu$ M for the endoA2 $\Delta$ H0 mutant). The force measurements were then repeated.

Pulling-force induced fission. After a scaffold has formed, confirmed with the drop in force, the source of the protein was removed and the tether was further elongated at a rate of 0.5  $\mu$ m s<sup>-1</sup>.

The radius of the tube was determined from the intensity of the fluorescence signal of the tube, considering that the ratio of fluorescence intensities of lipids inside the tube and inside the vesicle is proportional to the surface of the tube. To get absolute values, we used a normalization factor determined previously<sup>44</sup>.

**Image quantifications.** All image quantifications were performed with ImageJ or Fiji (NIH) and Matlab (MathWorks).

**Quantification of STxB transport to the Golgi apparatus.** Z-stacks were acquired on images of cells in defined experimental conditions. STxB-Cy3 fluorescence intensities were measured with ImageJ software (NIH) on z-projections, either from the entire cell, or from the Golgi region, as defined by giantin labelling. The ratio was then calculated as an index of Golgi localization.

**Quantification of transferrin-Alexa488 endocytosis.** The mean intensity of transferrin-Alexa488 fluorescence in the cell area was measured with ImageJ software (NIH) as above on z-projections, after background correction. Mean fluorescence intensities of control conditions were considered as 100%.

**Quantification of colocalization on TIRF images.** To quantify the colocalization between two channels, an object-based method was used, as implemented in JACoP<sup>45</sup>, based on the coincidence between two centroids with a 0.5 pixel tolerance. This was achieved in an ImageJ macro by first segmenting the tagged proteins by spot detection in each channel, finding their position, and growing them by dilation to 0.5 pixel radius. The spot detection consisted in finding maxima on the smoothed image (3  $\times$  3 average filter) using the 'find maxima' plugin of ImageJ, whose noise tolerance parameter was set up visually independently for each channel. The results were expressed as the percentage of colocalized spots over the total number of spots in the red and the green channel, respectively.

**Quantification of endoA2-GFP recruitment to the plasma membrane.** Spinning disk time series at fixed z-positions were acquired for 80 s at 37 °C with 0.84 s intervals and an exposure time of 80 ms for each channel. STxB was added 15 s after the beginning of each time series. Images were then treated with Fiji. Plasma membrane segments were selected randomly and linear regions of interest (ROI) were drawn. The mean fluorescence along ROI was measured in each channel over all images of a time series. Data were plotted for fluorescence intensity (in percentage of maximum intensity) over time (in seconds).

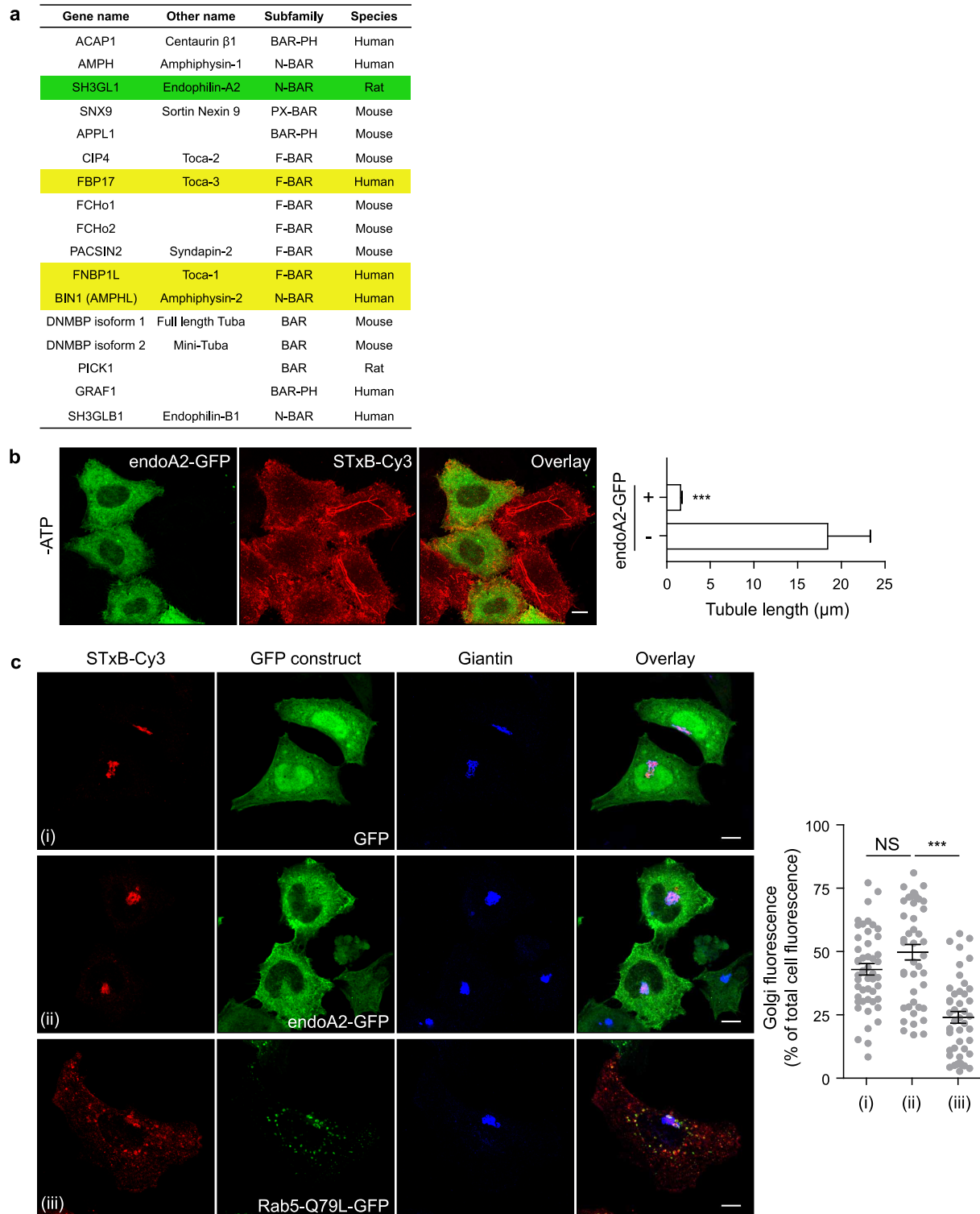
**Quantification of endoA2 spot lifetime.** Kymographs were obtained from image stacks using MetaMorph Software. Kymographs were processed using a first Fiji macro. The first step consisted in the creation of background images that were subtracted from the original images. Background images were obtained using the 'remove outliers' plugin with a 5-pixel radius and a 50 grey level threshold. An automatic threshold was then proposed and adjusted visually to select the tracks. Selected tracks were saved in ROI files for each image. A second macro was then used to measure the length of tracks, with a step allowing manual correction of ROI, if required. The lifetime of each structure was extracted from the length of the bounding boxes of the segmented tracks in the kymographs, corresponding to their length in frames.

**Quantification of STxB-induced tubule length.** The quantification of STxB-induced tubule length was achieved in batch for all images by segmenting the tubular structures and skeletonizing them so that the number of pixels of tubules was proportional to their length. For this, we used a Fiji macro based on the enhancement of tubular structures by computing eigenvalues of the hessian matrix on a Gaussian filtered version of the image (sigma = 1 pixel), as implemented in the tubeness plugin. The tubules were then thresholded and structures containing less than 3 pixels were discarded. A visual check was performed and a manual correction of segmented tubules was done, if required. The segmented and corrected structures were then reduced to a 1-pixel thick skeleton, using the Fiji plugin 'skeletonize'. Data were grouped by conditions and the length of skeletonized tubules in pixels was converted in micrometer units using Matlab (MathWorks).



**Statistical analyses.** All statistical analyses were performed using Prism v4.0 software (Graphpad Inc). Data were tested for Gaussian distribution with Kolmogorov–Smirnov test (with Dallal–Wilkinson–Lillie for *P* value). In case of non-Gaussian distribution, the following non-parametric tests were used: two-tailed Mann–Whitney *U* test if there were only two conditions or one-way ANOVA (Kruskal–Wallis test) with a Dunn’s test if there were more than two data groups to compare. In case of Gaussian distribution, the following parametric tests were used: two-tailed *t*-test for the comparison of the means if there were only two conditions to compare, parametric one-way ANOVA with a Bonferroni test if there were more than two data groups. Significance of mean comparison is represented on the graphs by asterisks. All error bars denote s.e.m., except if stated otherwise. No statistical method was used to predetermine sample size.

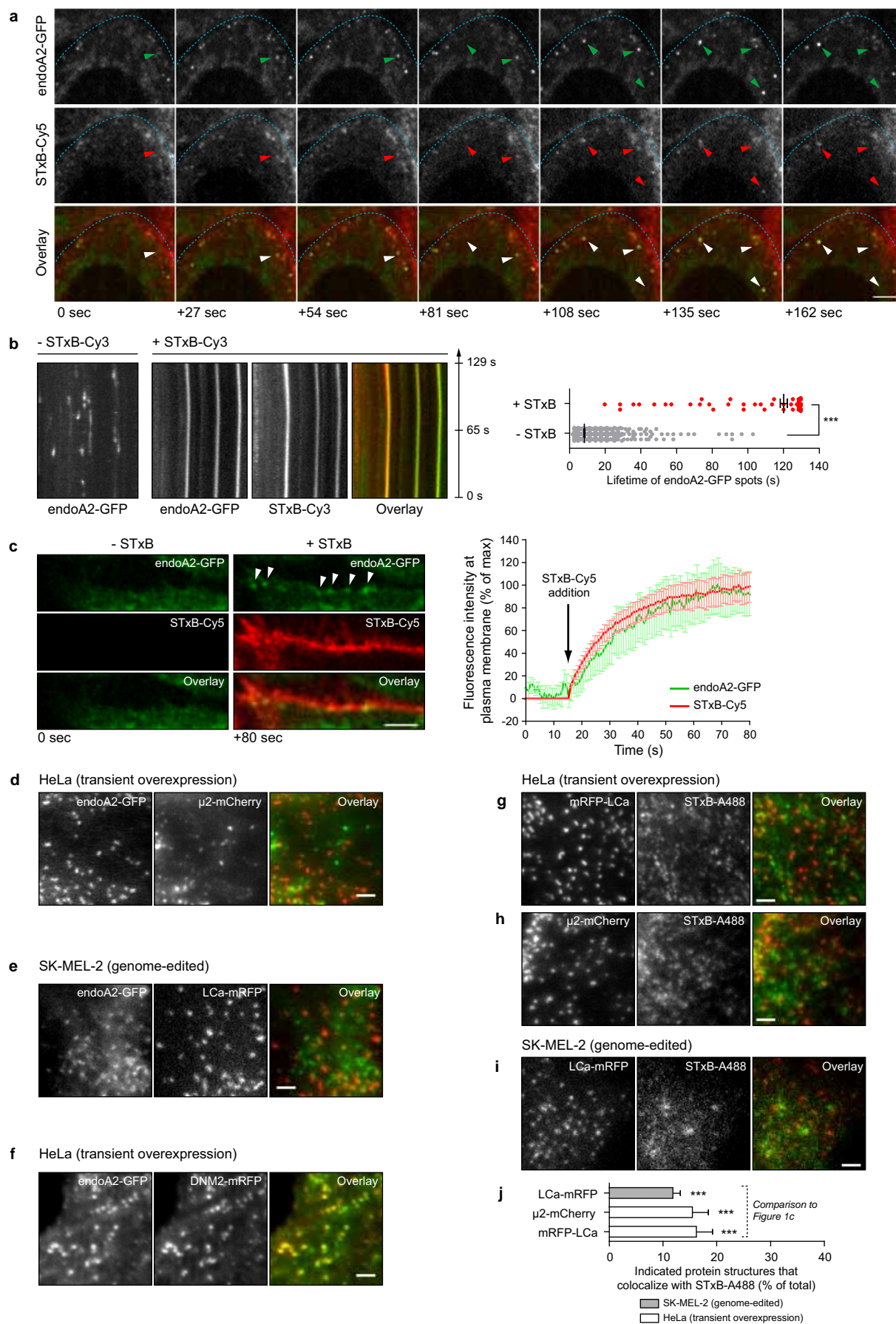
29. Zha, X. *et al.* Sphingomyelinase treatment induces ATP-independent endocytosis. *J. Cell Biol.* **140**, 39–47 (1998).
30. Lizárraga, F. *et al.* Diaphanous-related formins are required for invadopodia formation and invasion of breast tumor cells. *Cancer Res.* **69**, 2792–2800 (2009).
31. Riedl, J. *et al.* Lifeact: a versatile marker to visualize F-actin. *Nature Methods* **5**, 605–607 (2008).
32. Jordan, M., Schallhorn, A. & Wurm, F. M. Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res.* **24**, 596–601 (1996).
33. Montagnac, G. *et al.* Decoupling of activation and effector binding underlies ARF6 priming of fast endocytic recycling. *Curr. Biol.* **21**, 574–579 (2011).
34. Mallard, F. & Johannes, L. in *Shiga Toxin Methods and Protocols* Vol. 73 (eds Philpott, D. & Ebel, F.) Ch. 17, 209–220 (Humana Press, 2003).
35. Gortat, A., San-Roman, M. J., Vannier, C. & Schmidt, A. A. Single point mutation in Bin/Amphiphysin/Rvs (BAR) sequence of endophilin impairs dimerization, membrane shaping, and Src homology 3 domain-mediated partnership. *J. Biol. Chem.* **287**, 4232–4247 (2012).
36. Amessou, M., Popoff, V., Yelamos, B., Saint-Pol, A. & Johannes, L. Measuring retrograde transport to the *trans*-Golgi network. *Curr. Protoc. Cell. Biol.* **Chapter 15**, Unit–15.10 (2006).
37. Singh, R. D. *et al.* Selective caveolin-1-dependent endocytosis of glycosphingolipids. *Mol. Biol. Cell* **14**, 3254–3265 (2003).
38. Wolf, A. A. *et al.* Ganglioside structure dictates signal transduction by cholera toxin and association with caveolae-like membrane domains in polarized epithelia. *J. Cell Biol.* **141**, 917–927 (1998).
39. Stechmann, B. *et al.* Inhibition of retrograde transport protects mice from lethal ricin challenges. *Cell* **141**, 231–242 (2010).
40. Robinson, M. S. & Hirst, J. Rapid inactivation of proteins by knocksideways. *Curr. Protoc. Cell Biol.* **61**, 15.20.1–15.20.7 (2013).
41. Boncompain, G. & Perez, F. Synchronizing protein transport in the secretory pathway. *Curr. Protoc. Cell Biol.* **Chapter 15**, Unit–15.19 (2012).
42. Pautot, S., Frisken, B. J. & Weitz, D. A. Engineering asymmetric vesicles. *Proc. Natl Acad. Sci. USA* **100**, 10718–10721 (2003).
43. Pontani, L. L. *et al.* Reconstitution of an actin cortex inside a liposome. *Biophys. J.* **96**, 192–198 (2009).
44. Sorre, B. *et al.* Nature of curvature coupling of amphiphysin with membranes depends on its bound density. *Proc. Natl Acad. Sci. USA* **109**, 173–178 (2012).
45. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).



# Extended Data Figure 1 | Screening of BAR domain protein library.

**a**, Listing of BAR domain proteins that were tested in our localization screen. Toca-1, Toca-3 and amphiphysin-2 (yellow underlay) colocalized with STxB on tubular structures (not shown). See below for endoA2 (green underlay). **b**, At variance with the expected phenotype, expression of GFP-tagged endoA2 (green) led to the disappearance of long STxB-induced plasma membrane invaginations (red). Tubule length was quantified in non-transfected control cells ( $n = 50$ ) and endoA2-GFP-expressing cells ( $n = 59$ ). Quantifications show mean  $\pm$  s.e.m. of two independent experiments. \*\*\* $P < 0.001$

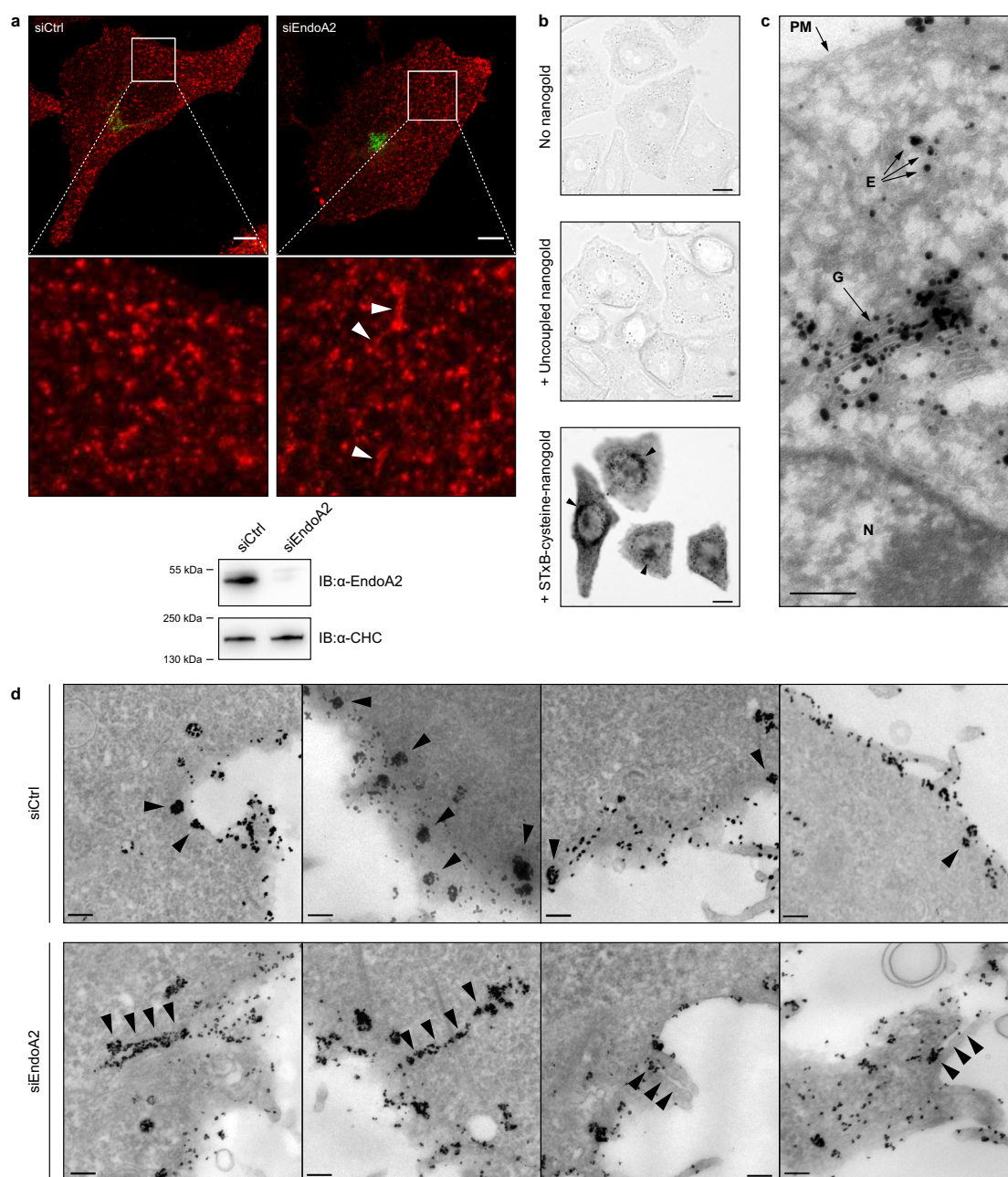
(two-tailed Mann-Whitney  $U$  test). **c**, STxB-Cy3 (50 nM, red) was incubated for 45 min at 37 °C with cells expressing GFP (i), endoA2-GFP (ii) or Rab5(Q79L)-GFP (iii). Expression of endoA2-GFP did not affect STxB trafficking to the Golgi/TGN membranes.  $n = 48$  for GFP-expressing cells,  $n = 43$  for endoA2-GFP-expressing cells, and  $n = 46$  for Rab5(Q79L)-GFP expressing cells; two independent experiments. Quantifications show mean  $\pm$  s.e.m. NS, non significant; \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). Scale bars, 10  $\mu$ m.



**Extended Data Figure 2 | Intracellular localization analysis of endoA2 and STxB.** **a–c**, STxB is internalized in endoA2–GFP-positive vesicles and induces the recruitment of endoA2–GFP to the plasma membrane. HeLaM cells stably expressing endoA2–GFP–FKBP (**a**, **c**), or HeLa cells transiently expressing endoA2–GFP (**b**) were incubated continuously at 37 °C respectively with 0.2  $\mu$ M STxB–Cy5 or 0.5  $\mu$ M STxB–Cy3. Observation by live-cell imaging using a spinning disk microscope (**a**, **c**) or TIRFM (**b**). **a**, STxB endocytosis. Image series of 170 s (for the complete sequence, see Supplementary Video 1). Arrows show the formation in the cell periphery of STxB–Cy5 and endoA2–GFP-positive vesicles that move into cells. **b**, Kymographs of HeLa cells transiently expressing endoA2–GFP in the absence (–) or presence (+) of STxB–Cy3. Quantification from TIRFM recordings of endoA2–GFP spot lifetime (–STxB,  $n = 1,768$  events; +STxB,  $n = 144$  events; three independent experiments. \*\*\* $P < 0.001$  (two-tailed Mann–Whitney  $U$  test). **c**, Plasma membrane recruitment of endoA2. STxB–Cy5 was added 15 s after the beginning of image acquisition. Two time points are shown (for the complete sequence, see Supplementary Video 2): before STxB addition (–STxB, 0 s) and after STxB addition (+STxB, +80 s). A striking recruitment of endoA2–GFP to the plasma membrane was observed (white arrowheads). Fluorescence

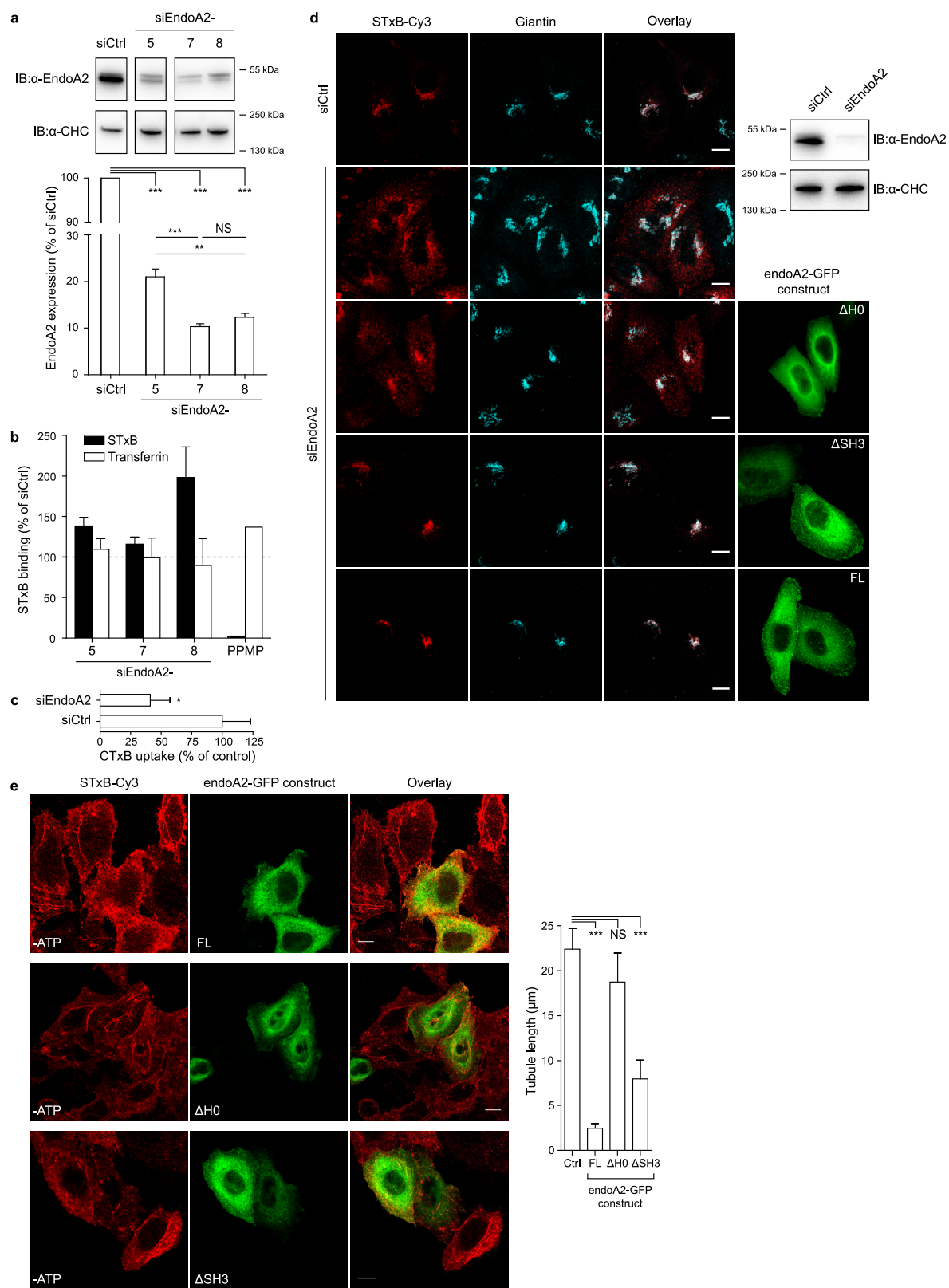
intensities along plasma membrane segments of endoA2–GFP and STxB–Cy5 were followed over time (means  $\pm$  s.e.m.,  $n = 6$  cells, three independent experiments). Note that both rose in a similar manner. **d–i**, EndoA2 and Shiga toxin poorly colocalize with markers of clathrin-mediated endocytosis. All images show live-cell TIRFM recordings. **d**, HeLa cells were transfected with plasmids expressing endoA2–GFP and  $\mu$ 2–mCherry. The overlap between both markers was very small. **e**, The genome-edited cell line SK-MEL-2 expressing LCa–mRFP was transfected transiently with the endoA2–GFP plasmid. Again, both markers showed very little overlap. **f**, HeLa cells were transfected with plasmids expressing endoA2–GFP and DNM2–mRFP. A substantial overlap was observed between both markers. For **d–f**, quantifications are reported in Fig. 1d. **g–i**, HeLa cells transiently expressing mRFP–LCa (**g**) or  $\mu$ 2–mCherry (**h**), and the genome-edited cell line SK-MEL-2 expressing LCa–mRFP (**i**) were continuously incubated with 0.5  $\mu$ M STxB–Alexa488 for 5 min at 37 °C. Note the weak overlap between STxB and the other markers. **j**, Quantification of colocalization for **g–i**. Mean  $\pm$  s.e.m. of the following numbers of cells: mRFP–LCa,  $n = 11$ ;  $\mu$ 2–mCherry,  $n = 6$ ; LCa–mRFP,  $n = 19$ ; two independent experiments. \*\*\* $P < 0.001$  (Bonferroni's multiple comparison test). Scale bars, 2  $\mu$ m.





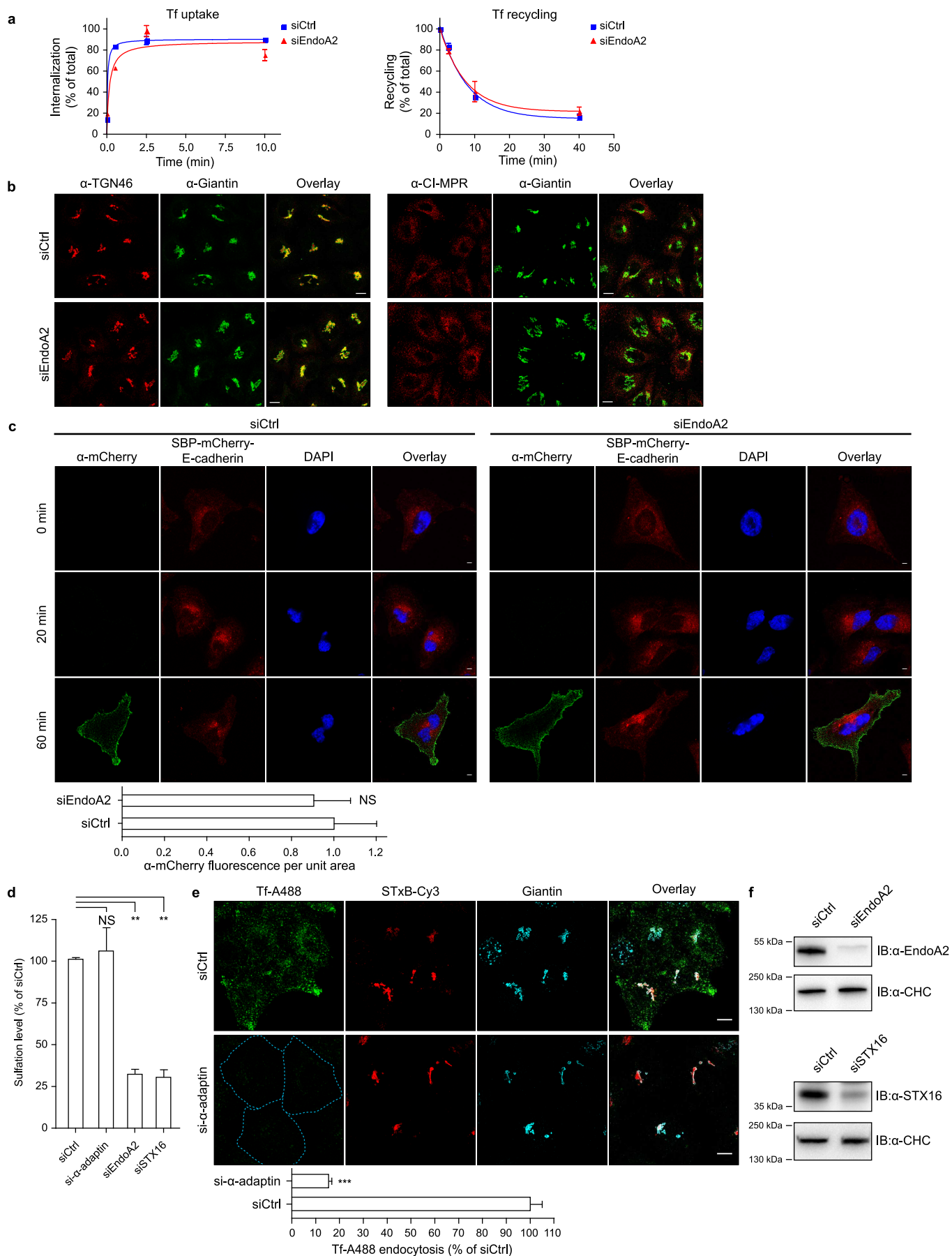
**Extended Data Figure 3 | Effect of endoA2 depletion on Shiga toxin endocytosis at early times of uptake.** **a**, After binding on ice for 30 min, STxB-Cy3 (50 nM, red) was incubated for 5 min at 37 °C with HeLa cells that had been transfected with negative control siRNAs (18 images), or siRNAs against endoA2 (23 images). After fixation, Golgi membranes were labelled with antibodies against giantin (green). Note the presence of short STxB-containing tubules (arrowheads) in the endoA2-depletion condition. The western blot documents the efficiency of endoA2 depletion. Clathrin heavy chain (CHC) was used as a loading control. **b**, **c**, Characterization of STxB conjugates with monofunctionalized nanogold. **b**, HeLa cells were incubated for 45 min at 37 °C in the indicated conditions, fixed, treated for silver enhancement, and viewed by transmission light microscopy. Note the strong perinuclear signal (arrowheads) that was visible only when cells were incubated

with nanogold-coupled STxB (113 images). **c**, HeLa cells were incubated for 45 min at 37 °C with nanogold-coupled STxB, fixed, treated for silver enhancement, and viewed by electron microscopy after sectioning (32 images). Note the strong signal in the Golgi region (G), which validated the functionality of the conjugate. Some STxB-nanogold could still be seen at the plasma membrane (PM) and in endosomes (E). N, nucleus. **d**, HeLa cells were transfected with negative control siRNAs, or siRNAs against endoA2, incubated for 30 min on ice with nanogold-coupled STxB, shifted for 5 min to 37 °C, and then fixed. Silver enhancement was used to enlarge nanogold particles. Note that STxB-containing invaginations (arrowheads) were much longer in endoA2-depleted cells than in cells transfected with negative control siRNA. Quantification is shown in Fig. 2b (see legend for cell numbers). Scale bars, 10  $\mu$ m (**a**, **b**), 200 nm (**c**), 100 nm (**d**).



**Extended Data Figure 4 | Effect of endoA2 depletion on STxB and CTxB uptake, and rescue with endoA2 mutants.** **a**, Efficiency of endoA2 depletion with different siRNA sequences. HeLa cells were transfected with three different siRNA sequences against human endoA2 (5, 7 and 8). The efficiency of endoA2 depletion was monitored by western blotting with antibodies against endoA2. Cells transfected with negative control siRNA were used for comparison. The western signal for CHC served as loading control. Quantification shows mean  $\pm$  s.e.m.; three independent experiments. NS, non significant;  $**P < 0.01$ ,  $***P < 0.001$  (Bonferroni's multiple comparison test). Trafficking studies on these cells are reported in Fig. 2c. **b**, Effect of endoA2 depletion on STxB binding to cells. HeLa cells were transfected with siRNAs as in **a**, detached, incubated for 30 min on ice with 1  $\mu$ M STxB-Alexa488 and 10  $\mu$ g ml<sup>-1</sup> Tf-Alexa647, washed and analysed by FACS. Cells treated with the glycosylceramide synthase inhibitor PPMP were used as a control for signal specificity. Mean  $\pm$  s.e.m. of three independent experiments are shown (except for PPMP, which was done twice). The increased STxB binding with siRNA 8 is not explained at this stage. These cell surface binding data serve as controls for the sulfation experiment of Fig. 2c. **c**, EndoA2 functions in CTxB uptake. CTxB (5 nM) uptake assay (three independent experiments) after 10 min at 37 °C in conditions of endoA2 depletion.  $*P < 0.05$  (two-tailed *t*-test). As opposed to STxB, CTxB could be removed from the plasma

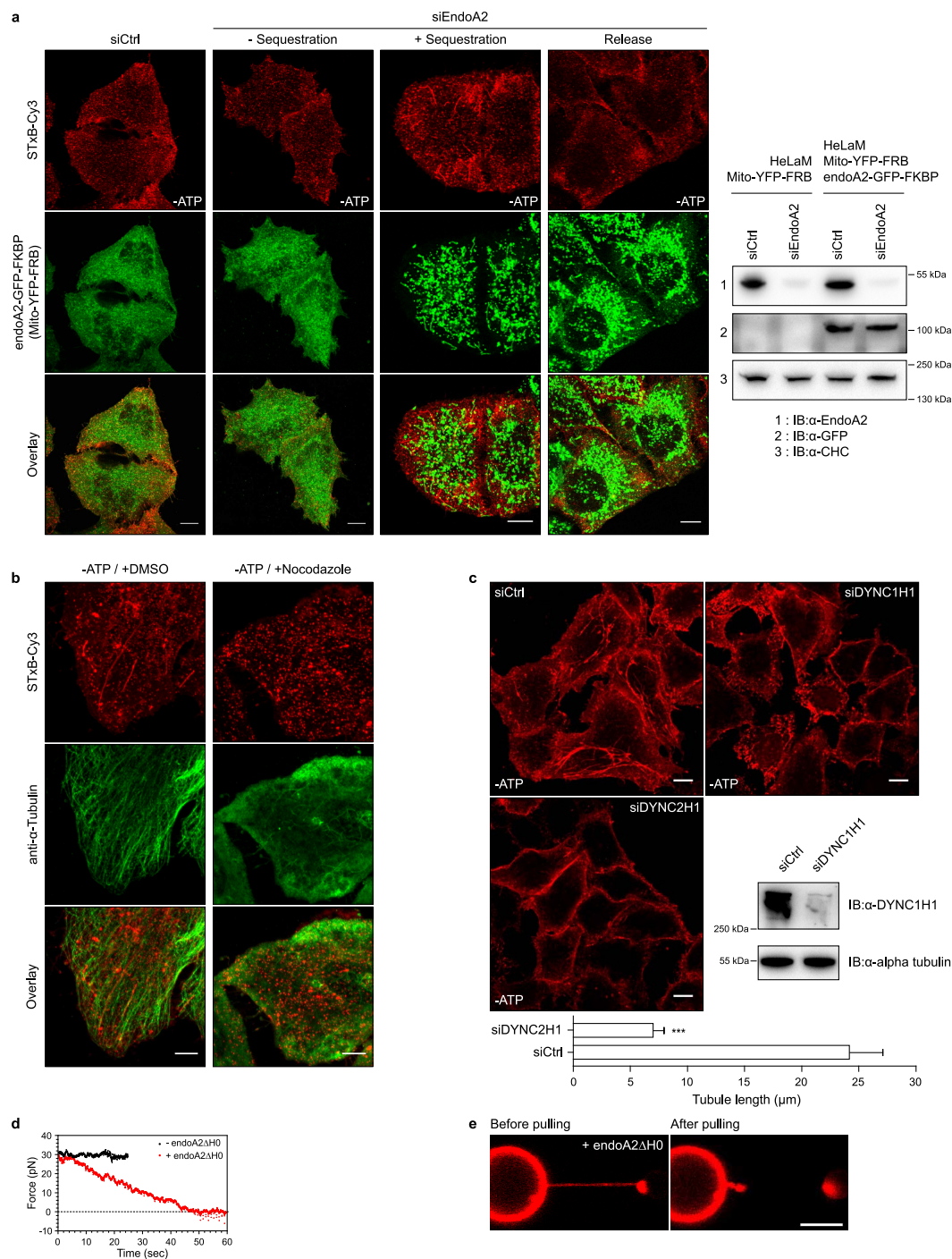
membrane by acid washes, and endocytosis could therefore be measured directly using a plate-reader assay (see Supplementary Methods). **d**, STxB trafficking rescue experiment under endoA2-depletion conditions. HeLa cells were transfected with negative control siRNAs, or siRNAs against endoA2. After 48 h, endoA2-depleted cells were transfected for 24 h with siRNA-resistant expression vectors encoding: GFP-tagged H0 helix deletion mutant (endoA2 $\Delta$ H0-GFP or  $\Delta$ H0), SH3 domain deletion mutant (endoA2 $\Delta$ SH3-GFP or  $\Delta$ SH3), or full-length GFP-tagged wild-type endoA2 (FL). STxB-Cy3 (50 nM, red) was incubated with these cells for 45 min at 37 °C (for quantification and statistical data, see Fig. 2d). The H0 helix deletion mutant did not rescue STxB transport to perinuclear Golgi/TGN membranes, as opposed to wild-type endoA2 and endoA2 $\Delta$ SH3. The western blot documents the efficiency of endoA2 depletion in siRNA-transfected cells. CHC was used as a loading control. **e**, Effect of endoA2 mutants on STxB-induced membrane invaginations. Non-transfected HeLa cells (ctrl) or cells expressing full-length endoA2 (FL), H0 helix deletion mutant ( $\Delta$ H0), or SH3 domain deletion mutant ( $\Delta$ SH3) were ATP-depleted and incubated for 10 min at 37 °C with 0.5  $\mu$ M STxB-Cy3. Tubule length was quantified (control,  $n = 277$  in 102 cells; full-length,  $n = 90$  in 15 cells;  $\Delta$ H0,  $n = 183$  in 48 cells;  $\Delta$ SH3,  $n = 164$  in 36 cells); two independent experiments. Quantifications show mean  $\pm$  s.e.m.  $***P < 0.001$  (Bonferroni's multiple comparison test). Scale bars, 10  $\mu$ m.





**Extended Data Figure 5 | Intracellular trafficking analysis.** **a–c**, EndoA2 depletion does not affect general trafficking processes. All experiments were performed on cells transfected with negative control siRNAs, and cells that were depleted for endoA2. **a**, Transferrin uptake (left,  $n = 3$ ) and recycling (right,  $n = 4$ ). EndoA2 depletion did not affect any of these processes. Quantifications show mean  $\pm$  s.e.m. of the indicated numbers of independent experiments. **b**, Steady-state localization of TGN46 and CI-MPR (red), as analysed by immunofluorescence. Golgi membranes were labelled with antibodies against giantin (green). EndoA2 depletion did not affect the steady-state localization of these markers (TGN46 in siCtrl or siEndoA2 cells: 12 images; CI-MPR in siCtrl or siEndoA2 cells: 10 images; two independent experiments). **c**, Anterograde trafficking of E-cadherin. After release from endoplasmic reticulum, SBP-mCherry-E-cadherin protein was detected at the cell surface with anti-mCherry antibodies (0, 20 and 60 min time points). After 60 min, the relative means ( $\pm$  s.e.m.) of anti-mCherry fluorescence per unit area were quantified for control (6 images, 87 cells) and endoA2-depleted cells (6 images, 81 cells); three independent experiments. No significant difference was observed using a two-tailed  $t$ -test ( $P > 0.05$ ). **d–f**, Depletion of  $\alpha$ -adaptin does not affect Shiga toxin trafficking. All experiments were performed on negative control siRNA-transfected cells, and on cells that were

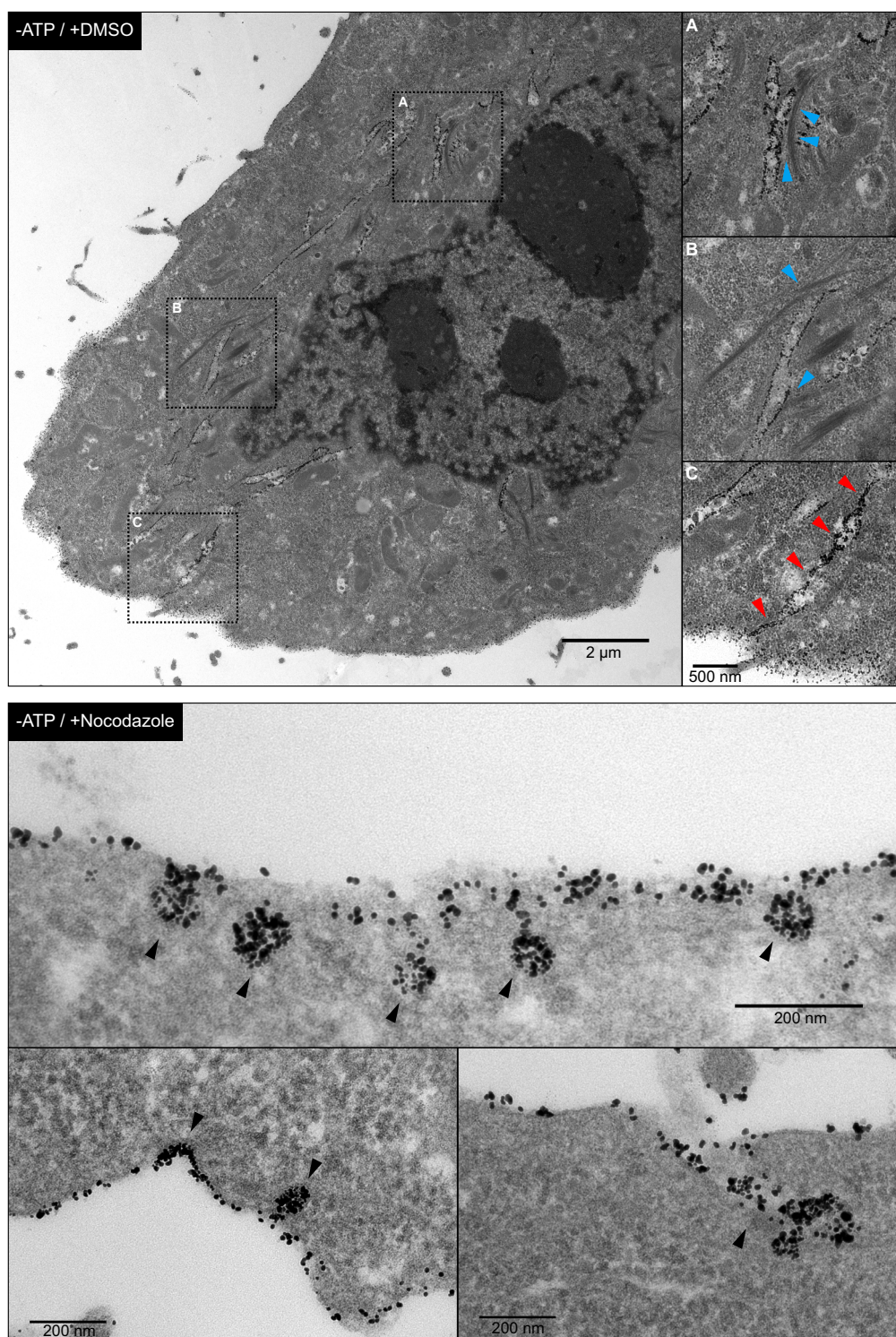
depleted for the indicated proteins. Quantifications show mean  $\pm$  s.e.m. **d**, Sulfation analysis of retrograde STxB transport (three independent experiments). HeLa cells in the indicated conditions were incubated for 20 min at 37 °C with STxB-Sulf<sub>2</sub> in the presence of radioactive sulphate, and sulfated STxB-Sulf<sub>2</sub> was measured by autoradiography. Note that  $\alpha$ -adaptin depletion did not affect sulfation of STxB-Sulf<sub>2</sub>, whereas depletion of endoA2 or syntaxin-16 had a strong effect. NS, non significant; \*\* $P < 0.01$  (Bonferroni's multiple comparison test). **e**, Immunofluorescence analysis. HeLa cells in the indicated conditions were incubated for 45 min at 37 °C with 0.05  $\mu$ M STxB-Cy3 (red). During the last 10 min, 10  $\mu$ g ml<sup>-1</sup> Tf-Alexa488 (green) were added in the growth medium. Cells were placed on ice, and cell-surface-exposed transferrin was removed by acid washes. After fixation, cells were labelled for giantin (blue). Note that  $\alpha$ -adaptin depletion strongly inhibited transferrin uptake, but not retrograde transport of STxB to TGN/Golgi membranes. Transferrin uptake was quantified for control (5 images, 102 cells) and  $\alpha$ -adaptin-depleted cells (5 images, 108 cells); two independent experiments. \*\*\* $P < 0.001$  (two-tailed  $t$ -test). **f**, siRNA-mediated depletion of endoA2 and of syntaxin-16 was analysed by western blotting. CHC was used as loading control. Scale bars, 10  $\mu$ m.



### Extended Data Figure 6 | Cell and model membrane experiments.

**a**, Knocksideways. HeLaM cells stably expressing Mito-YFP-FRB and rat endoA2-GFP-FKBP (green) were transfected with negative control siRNAs (siCtrl), or siRNAs against human endoA2 (siEndoA2) that did not cross with the rat sequence. STxB-Cy3 (0.5 μM, red) was incubated for 15 min at 37 °C with ATP-depleted cells. The cells were then fixed at 37 °C, and viewed by confocal microscopy. Quantification of tubule formation and cell numbers are shown in Fig. 3a. Note that STxB-induced tubule length reversibly increased after endoA2-GFP-FKBP sequestration. The depletion of endogenous human endoA2, and the expression of GFP-FKBP-tagged rat endoA2 were assessed by western blotting with anti-endoA2 and anti-GFP antibodies, respectively. Western blotting against CHC was used as loading control. **b**, **c**, Interfering with microtubules or dynein motors strongly affects STxB-induced tubule length. **b**, HeLa cells were incubated for 1 h at 37 °C with DMSO or 10 μM nocodazole, ATP-depleted for 20 min, and then incubated for 10 min at 37 °C with 0.5 μM STxB-Cy3 (red) in the same conditions. Labelling with an antibody against α-tubulin (green) was used to visualize the efficiency of

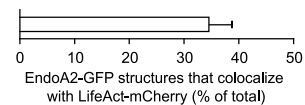
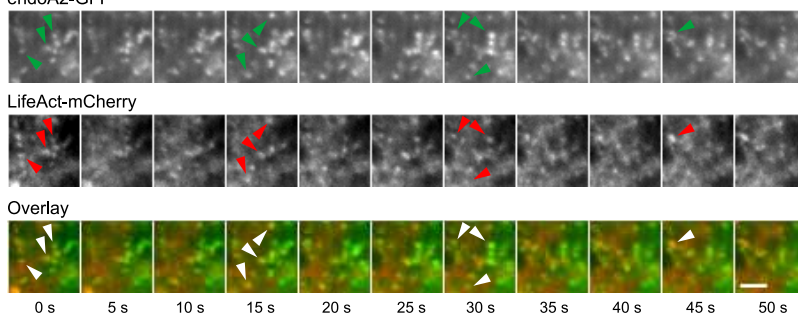
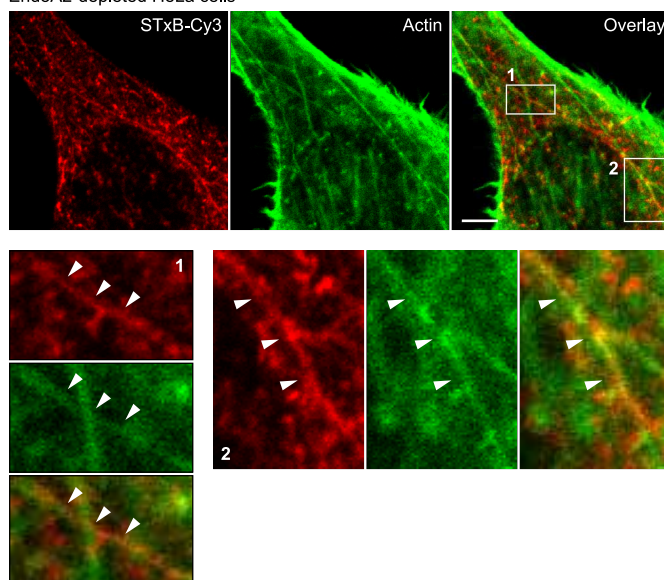
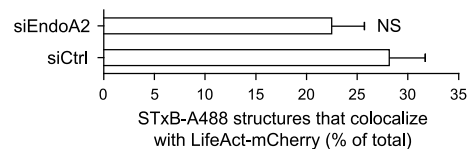
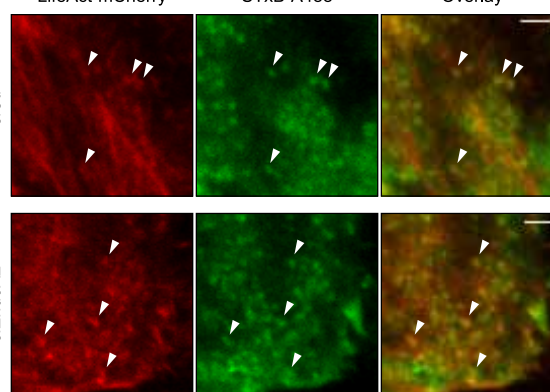
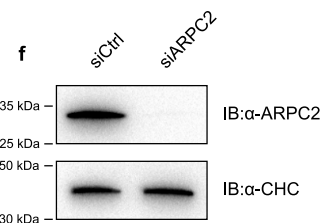
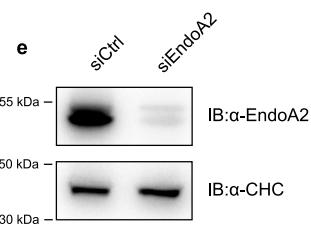
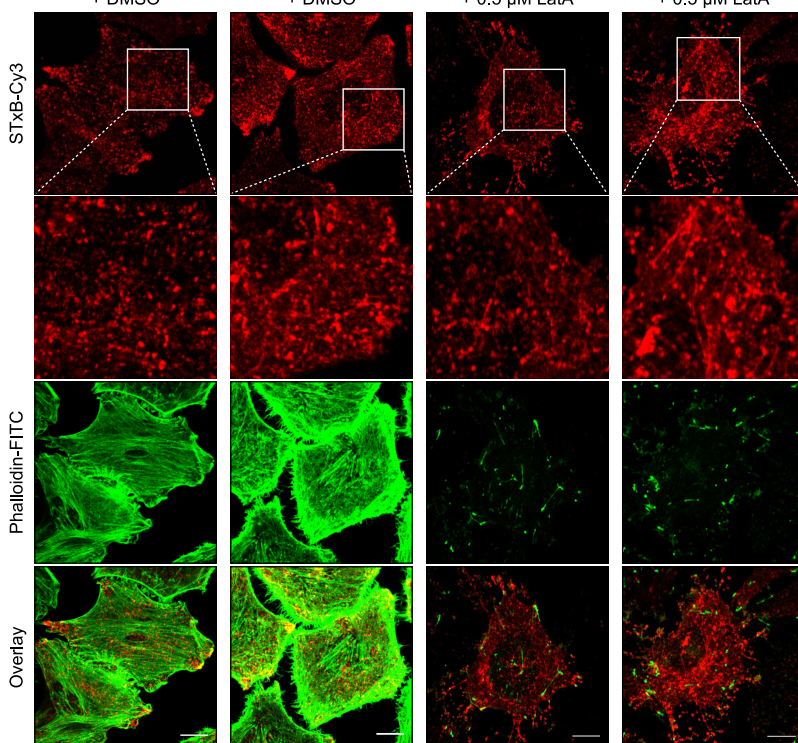
nocodazole treatment. Long tubular structures containing STxB could not be detected after incubation with nocodazole (−ATP/+DMSO: 18 images; −ATP/+nocodazole: 16 images; three independent experiments). **c**, Heavy chains of cytoplasmic dyneins (DYNC1H1 and DYNC2H1) were depleted from HeLa cells with siRNAs. Cells were ATP-depleted and then incubated for 10 min at 37 °C with 0.5 μM STxB-Cy3 (red). The presence of long STxB-induced tubules was strongly decreased under these conditions. Tubule length was quantified for negative-control-siRNA-treated cells (siCtrl,  $n = 188$  tubules in 74 cells) and for DYNC2H1-depleted cells (siDYNC2H1,  $n = 165$  tubules in 60 cells); two independent experiments. Quantifications show means ± SEM. \*\*\* $P < 0.001$ , two-tailed Mann-Whitney  $U$  test. The western blot documents the efficiency of DYNC1H1 depletion. α-tubulin was used as a loading control. **d**, **e**, Model membranes experiments. **d**, Measurement of tube pulling force over time in the absence or presence of endoA2ΔH0 mutant (7 μM in injection pipette). Representative of six experiments. **e**, Scission experiments with tethers that were coated with endoA2ΔH0 ( $n = 5$ ). Scale bars, 10 μm (**a**, **c**) and 5 μm (**b**, **e**).



**Extended Data Figure 7 | Electron microscopy of STxB-nanogold on ATP-depleted cells following treatment with nocodazole.** HeLa cells were treated for 1 h at 37 °C with DMSO (top) or 10 µM nocodazole (bottom). During the last 20 min, ATP was depleted. Cells were then incubated for 10 min with STxB-nanogold in the continued presence of inhibitors, fixed, and prepared for electron microscopy. In the DMSO condition (top), long tubular structures connected to the cell surface were observed (magnified views in right insets, red arrowheads in C), as expected from light microscopy

experiments. These structures were in close proximity with cytoskeletal elements, as indicated with blue arrowheads in magnifications A and B. In the nocodazole condition (bottom), STxB-induced plasma membrane invaginations were still present (arrowheads), but much shorter (mean length of  $118.2 \pm 7.0$  nm,  $n = 109$  invaginations;  $0.90 \pm 0.12$  invaginations/µm of plasma membrane,  $n = 28$  images; three independent experiments) than in the absence of the compound. Scale bar sizes are indicated.

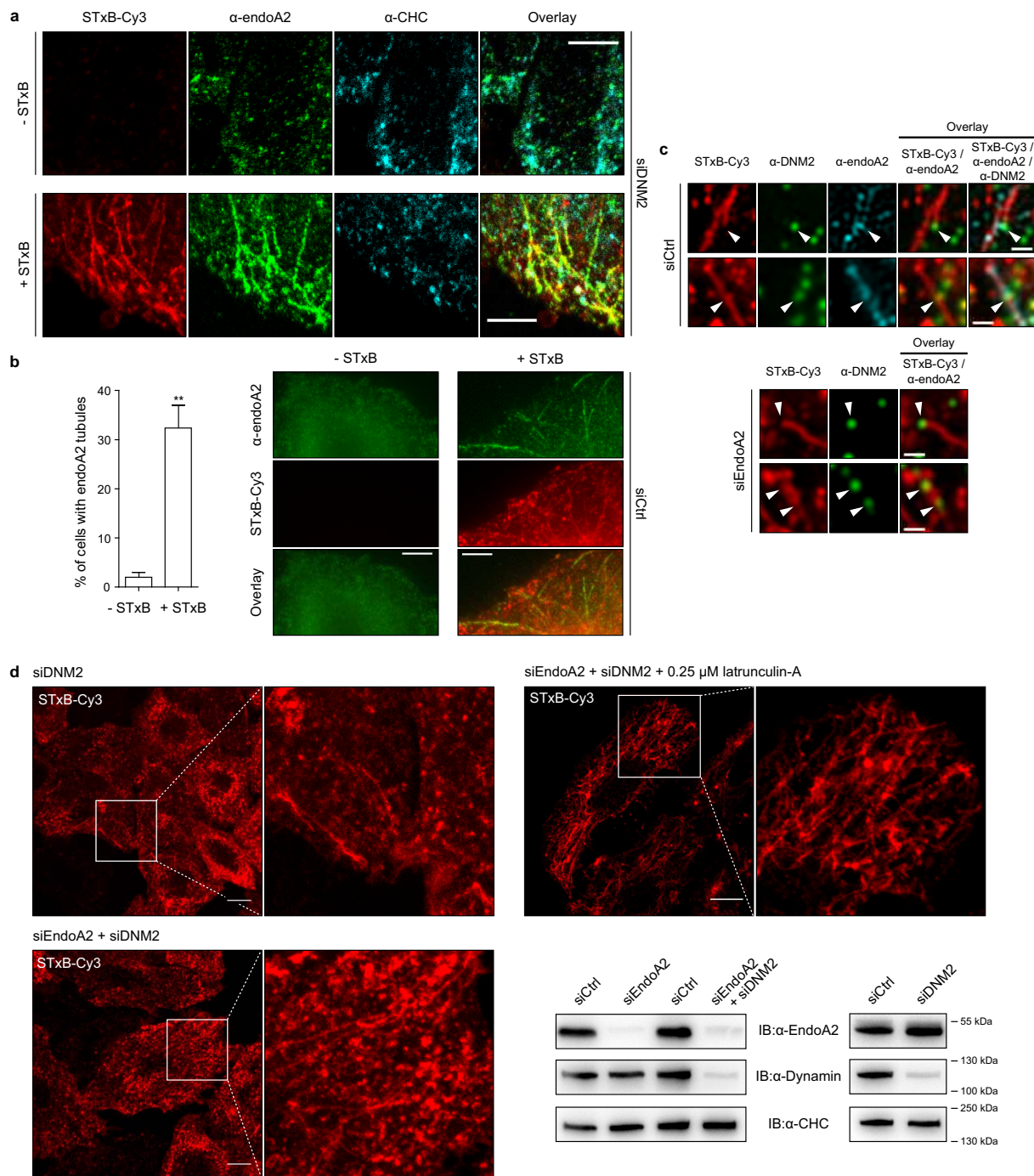


**a** endoA2-GFP**b** EndoA2-depleted HeLa cells**c** LifeAct-mCherry STxB-A488 Overlay**d** siCtrl + DMSO siEndoA2 + DMSO siCtrl + 0.5  $\mu$ M LatA siEndoA2 + 0.5  $\mu$ M LatA



**Extended Data Figure 8 | Actin and endoA2.** **a–c**, EndoA2 codistribution with actin. **a**, HeLa cells transiently co-expressing endoA2-GFP and LifeAct-mCherry were observed by time-resolved TIRF microscopy. The panel shows acquisitions at the plasma membrane that were taken at 5-s intervals. Arrowheads point out examples of structures on which endoA2 and actin colocalize in a dynamic manner. Quantification of colocalization of endoA2-positive structures with LifeAct is presented as mean  $\pm$  s.e.m. ( $n = 7$  cells, 2 independent experiments). **b**, EndoA2-depleted HeLa cells were incubated continuously for 5 min at 37 °C with 0.5  $\mu$ M STxB-Cy3. After fixation, actin filaments were stained with phalloidin-FITC. Arrowheads indicate STxB-induced tubules that are decorated by actin. **c**, HeLa cells transfected with negative control or endoA2 siRNAs and transiently expressing LifeAct-mCherry were observed by TIRF microscopy after addition of 0.5  $\mu$ M STxB-Cy3 at 37 °C. Arrowheads point out examples of structures on which STxB and actin colocalize. Quantification of colocalization of STxB-positive structures with LifeAct is presented as means  $\pm$  s.e.m. ( $n = 6$  cells, two independent

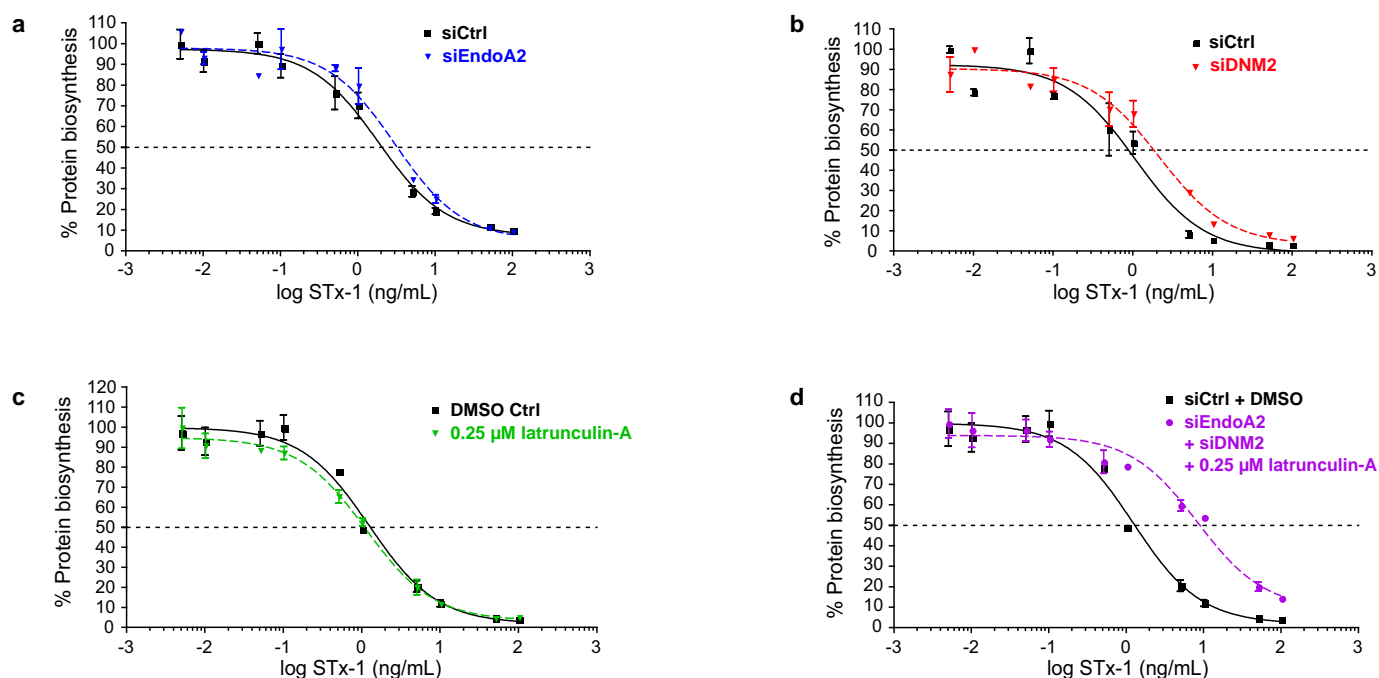
experiments). NS, non significant (two-tailed  $t$ -test). **d, e**, Analysis of STxB-induced plasma membrane invaginations in function of endoA2 depletion and/or actin perturbation. **d**, HeLa cells were transfected with negative control siRNAs or with siRNAs against endoA2, and treated or not with 0.5  $\mu$ M latrunculin-A. The cells were then incubated continuously for 5 min at 37 °C in the presence of 0.5  $\mu$ M STxB-Cy3 (red), fixed and labelled with phalloidin (green). The quantification of STxB tubule length is shown in Fig. 4a. Note that tubule length increased with combined treatments (siCtrl+DMSO: 22 images; siEndoA2+DMSO: 11 images; siCtrl+LatA: 14 images; siEndoA2+LatA: 14 images). **e**, Depletion of endoA2 was analysed by western blotting. CHC was used as loading control. **f**, Efficiency of ARPC2 depletion. HeLa cells were transfected for 72 h with a smartpool of four siRNA sequences against ARPC2. The efficiency of ARPC2 depletion was monitored by western blotting with antibodies against ARPC2. The western signal for CHC served as loading control. Corresponds to experiments in Fig. 4a. Scale bars, 2  $\mu$ m (**a, c**), 5  $\mu$ m (**b**) and 10  $\mu$ m (**d**).



### Extended Data Figure 9 | Combined effects of interference with endoA2, dynamin and actin on STxB-induced membrane invaginations.

**a–c**, Endogenous endoA2 and dynamin are found on STxB-induced plasma membrane invaginations. **a**, Dynamin-2 was depleted from cells (siDNM2), which were then incubated continuously for 5 min at 37 °C in the presence or absence of 0.5  $\mu$ M STxB–Cy3 (red), fixed, labelled for the indicated markers, and analysed by confocal microscopy. Note that endoA2-containing tubules (green) were seen only in the presence of STxB. No overlap was observed with clathrin (blue). –STxB: representative of 15 images; +STxB: representative of 35 images; two independent experiments. **b**, Experiment as in **a** on wild-type cells that were analysed by wide field microscopy. As above, endoA2 (green) was found on tubular structures only in the presence of STxB (red). The quantification shows mean  $\pm$  s.e.m. of 3 independent experiments on 234 cells without STxB (–STxB) and 921 cells with STxB (+STxB). \*\* $P$  < 0.01, two-tailed  $t$ -test. **c**, Negative control siRNA transfected HeLa cells and endoA2-depleted cells were incubated for 5 min at 37 °C in the presence of 0.5  $\mu$ M STxB–Cy3 (red). Endogenous endoA2 (blue) and dynamin

(green, arrowheads) were labelled with specific antibodies, detected by immunofluorescence, and viewed by structured illumination microscopy. Note that dynamin localized in spots on STxB-induced invaginations, while endoA2 distributed in a continuous manner. **d**, Analysis of STxB-induced plasma membrane invaginations in function of endoA2 and dynamin-2 depletion, and actin perturbation. HeLa cells were depleted for dynamin-2 alone, dynamin-2 in combination with endoA2, or both depletions in combination with 0.25  $\mu$ M latrunculin-A treatment, as indicated. These cells were then incubated continuously for 5 min at 37 °C with 0.5  $\mu$ M STxB–Cy3 (red), fixed at 37 °C, and viewed by confocal microscopy. Note that the tubulation phenotype increased with each additional interference modality (siDNM2: 21 images; siEndoA2 + siDNM2: 15 images; siEndoA2 + siDNM2 + LatA: 16 images; two independent experiments). The quantification of tubule length in the different experimental conditions of this figure is shown in Fig. 4a. The depletion of dynamin-2 and endoA2 was validated by immunoblotting. CHC was used as loading control. Scale bars, 5  $\mu$ m (**a**, **b**), 0.5  $\mu$ m (**c**) and 10  $\mu$ m (**d**).



**Extended Data Figure 10 | Intoxication curves.** a–d, HeLa cells were depleted for endoA2 (a), dynamin-2 (b), incubated with 0.25  $\mu$ M latrunculin-A (c), or submitted concomitantly to all three perturbations (d). These cells were then further incubated for 1 h in the presence of increasing concentrations of STx-1, at the end of which protein biosynthesis was measured. Note that only the

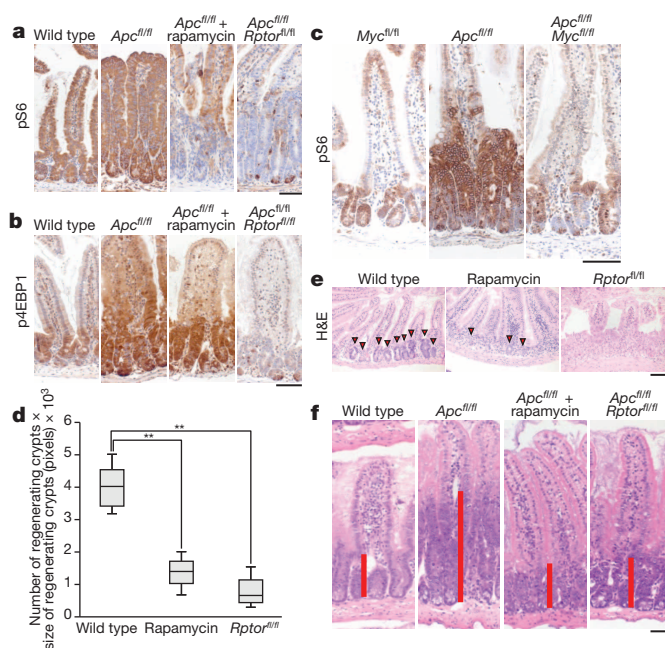
triple treatment condition had a strong effect on cell intoxication. The protection factors determined on 4–13 independent experiments are shown in Fig. 4c. siEndoA2,  $n = 7$ ; siDNM2,  $n = 4$ ; latrunculin-A,  $n = 13$ ; triple,  $n = 4$ . Error bars show the s.e.m.

# mTORC1-mediated translational elongation limits intestinal tumour initiation and growth

William J. Faller<sup>1</sup>, Thomas J. Jackson<sup>2\*</sup>, John R. P. Knight<sup>2\*</sup>, Rachel A. Ridgway<sup>1</sup>, Thomas Jamieson<sup>1</sup>, Saadia A. Karim<sup>1</sup>, Carolyn Jones<sup>2</sup>, Sorina Radulescu<sup>1</sup>, David J. Huels<sup>1</sup>, Kevin B. Myant<sup>1</sup>, Kate M. Dudek<sup>2</sup>, Helen A. Casey<sup>1</sup>, Alessandro Scopelliti<sup>1</sup>, Julia B. Cordero<sup>1</sup>, Marcos Vidal<sup>1</sup>, Mario Pende<sup>3</sup>, Alexey G. Ryazanov<sup>4</sup>, Nahum Sonenberg<sup>5</sup>, Oded Meyuhas<sup>6</sup>, Michael N. Hall<sup>7</sup>, Martin Bushell<sup>2</sup>, Anne E. Willis<sup>2</sup> & Owen J. Sansom<sup>1</sup>

Inactivation of APC is a strongly predisposing event in the development of colorectal cancer<sup>1,2</sup>, prompting the search for vulnerabilities specific to cells that have lost APC function. Signalling through the mTOR pathway is known to be required for epithelial cell proliferation and tumour growth<sup>3–5</sup>, and the current paradigm suggests that a critical function of mTOR activity is to upregulate translational initiation through phosphorylation of 4EBP1 (refs 6, 7). This model predicts that the mTOR inhibitor rapamycin, which does not efficiently inhibit 4EBP1 (ref. 8), would be ineffective in limiting cancer progression in APC-deficient lesions. Here we show in mice that mTOR complex 1 (mTORC1) activity is absolutely required for the proliferation of *Apc*-deficient (but not wild-type) enterocytes, revealing an unexpected opportunity for therapeutic intervention. Although APC-deficient cells show the expected increases in protein synthesis, our study reveals that it is translation elongation, and not initiation, which is the rate-limiting component. Mechanistically, mTORC1-mediated inhibition of eEF2 kinase is required for the proliferation of APC-deficient cells. Importantly, treatment of established APC-deficient adenomas with rapamycin (which can target eEF2 through the mTORC1–S6K–eEF2K axis) causes tumour cells to undergo growth arrest and differentiation. Taken together, our data suggest that inhibition of translation elongation using existing, clinically approved drugs, such as the rapalogs, would provide clear therapeutic benefit for patients at high risk of developing colorectal cancer.

The ability of the intestinal epithelium to regenerate after challenge has been well described<sup>9–11</sup>. We have shown that this is a Wnt-driven process that mimics the proliferation observed after *Apc* deletion<sup>11,12</sup> and is a valuable model of the early stages of intestinal cancer. However, the underlying mechanisms controlling these processes are largely unknown. The serine/threonine kinase mTOR, particularly as part of mTORC1, is a known mediator of cell growth and proliferation<sup>13</sup>. Previous studies have suggested that mTORC1 may be important in both the intestinal stem-cell niche and for intestinal tumorigenesis<sup>4,5,14</sup>. We therefore queried the role of mTORC1 in intestinal proliferation after Wnt activation. Following *Apc* deletion there was an increase in the phosphorylation status of the mTORC1 effectors RPS6 and 4EBP1 that was dependent on MYC expression. Increased phosphorylation of these proteins was also seen during crypt regeneration (Fig. 1a–c and Extended Data Fig. 1a). Importantly, the mTOR inhibitor rapamycin blocked intestinal regeneration, demonstrating that mTOR signalling is required for this process (Fig. 1d, e). Given that rapamycin did not affect apoptosis or proliferation in the normal intestine (Extended Data Fig. 1b, c), these data suggest that there may be a potential therapeutic window, between normal intestinal enterocytes and those with a high level of Wnt activity. Therefore, we deleted raptor (*Rptor*; an essential component of mTORC1) in



**Figure 1 | mTORC1 is essential for Wnt-driven proliferation in a MYC-dependent manner.** **a, b**, Representative immunohistochemistry (IHC) of phospho-RPS6 (pS6) and phospho-4EBP1 (p4EBP1) showing increased staining 96 h after *Apc* deletion. *Rptor* deletion caused a loss of positivity in both, whereas 10 mg kg<sup>-1</sup> rapamycin treatment (beginning at 24 h) specifically disrupts RPS6 phosphorylation (representative of six biological replicates). **c**, Representative IHC of phospho-RPS6 96 h after Cre induction showing that Wnt-driven RPS6 phosphorylation is MYC dependent (representative of 3 biological replicates). **d**, Mice were exposed to 14 Gy  $\gamma$ -irradiation and intestinal regeneration was measured 72 h later by counting the number of viable crypts and multiplying that by the average size of the regenerating crypts. Boxplot shows that 10 mg kg<sup>-1</sup> rapamycin treatment and *Rptor* deletion significantly decrease intestinal regeneration. Whiskers show maximum and minimum, black line shows median ( $n = 6$  biological replicates per group). \*\* $P$  value < 0.02, Mann–Whitney  $U$  test. **e**, Representative haematoxylin and eosin (H&E) staining of regenerating intestines 72 h after exposure to 14 Gy  $\gamma$ -irradiation. Arrowheads indicate regenerating crypts (representative of 6 biological replicates). **f**, Representative H&E staining 96 h after *Apc* loss, showing that 10 mg kg<sup>-1</sup> rapamycin treatment or *Rptor* deletion prevent Wnt-driven proliferation (representative of 6 biological replicates). Treatment began 24 h after *Apc* deletion. Red bar is graphical representation of crypt size. Scale bars, 100  $\mu$ m.

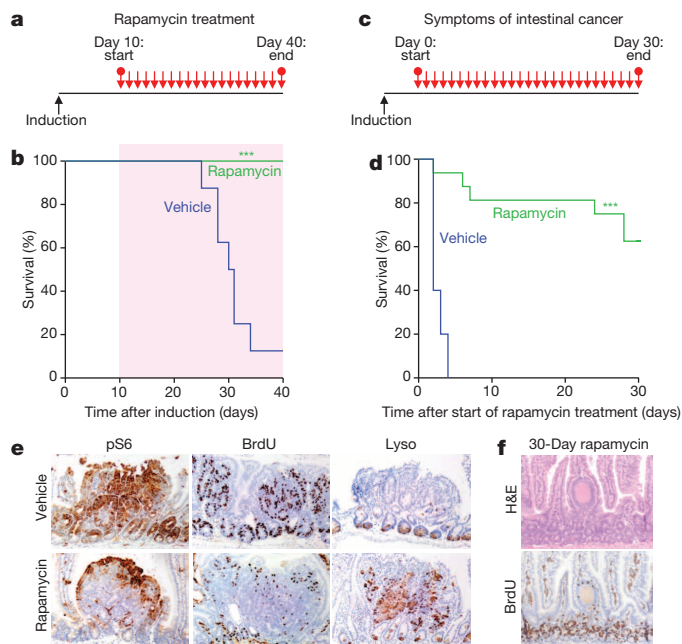
<sup>1</sup>Cancer Research UK Beatson Institute, Glasgow G61 1BD, UK. <sup>2</sup>Medical Research Council Toxicology Unit, Leicester LE1 9HN, UK. <sup>3</sup>Institut Necker-Enfants Malades, CS 61431, Paris, France Institut National de la Santé et de la Recherche Médicale, U1151, F-75014 Paris, France Université Paris Descartes, Sorbonne Paris Cité, 75006 Paris, France. <sup>4</sup>Department of Pharmacology, Rutgers The State University of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA. <sup>5</sup>Department of Biochemistry and Goodman Cancer Research Center, McGill University, Montreal, Québec H3A 1A3, Canada. <sup>6</sup>Department of Biochemistry and Molecular Biology, IMRIC, The Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel. <sup>7</sup>Biozentrum, University of Basel, CH-4056 Basel, Switzerland.

\*These authors contributed equally to this work.



the intestinal epithelium (Extended Data Fig. 1d). Surprisingly, normal gut homeostasis was unaffected by raptor loss 4 days after Cre induction, when using an epithelium-specific Cre-recombinase (*VillinCre<sup>ER</sup>Rptor<sup>fl/fl</sup>*) (Extended Data Fig. 1e, f). Furthermore, 400 days after induction, no phosphorylation of RPS6 or 4EBP1 was observed, showing that *Rptor* deletion was sustained (Extended Data Fig. 2a). Raptor loss caused no change in levels of either mitosis or apoptosis (Extended Data Fig. 2b, c) but proved to be essential for the proliferative phenotype observed during regeneration or after *Apc* deletion (Fig. 1a, b, d–f). Nuclear localization of  $\beta$ -catenin and high levels of MYC could be demonstrated by immunohistochemistry (IHC), showing that Wnt activation is still present (Extended Data Fig. 3a, b).

Given that rapamycin treatment and *Rptor* deletion had similar effects, we examined whether rapamycin treatment was sufficient to modify intestinal tumorigenesis, either prophylactically or chemotherapeutically. First we assessed whether rapamycin could suppress a model of intestinal tumorigenesis, in which *Apc* deletion is targeted to Lgr5-positive stem cells using *Lgr5Cre<sup>ER</sup> Apc<sup>fl/fl</sup>*. Mice were treated starting 10 days after Cre induction and, in contrast to controls, remained tumour free for the duration of the experiment (Fig. 2a, b). Next we treated mice (*Apc<sup>Min/+</sup>* or *LGR5Cre<sup>ER</sup> Apc<sup>fl/fl</sup>*) with established adenomas. Remarkably, the mice lost their clinical symptoms of disease and survived significantly



**Figure 2 | Tumorigenesis driven by the loss of *Apc* requires mTORC1 activation.** **a, b**, Graphical representation of prophylactic rapamycin treatment strategy and Kaplan–Meyer survival curve showing that prophylactic rapamycin treatment prevents tumorigenesis. Rapamycin treatment ( $10 \text{ mg kg}^{-1}$ ) began at day 10 after *Apc* deletion, and lasted 30 days, after which mice were sampled. Area highlighted by red indicates duration of rapamycin treatment ( $n = 8$ , vehicle;  $n = 13$ , rapamycin). \*\*\* $P$  value  $\leq 0.001$ , log-rank test. **c, d**, Graphical representation of chemotherapeutic rapamycin treatment strategy and Kaplan–Meyer survival curve showing that rapamycin treatment can regress established intestinal tumours. Rapamycin treatment ( $10 \text{ mg kg}^{-1}$ ) started when mice showed signs of intestinal disease, and lasted 30 days, after which mice were sampled. Graph represents survival while on rapamycin treatment ( $n = 5$ , vehicle;  $n = 16$ , rapamycin). \*\*\* $P$  value  $\leq 0.001$ , log-rank test. **e**, Representative IHC of phospho-RPS6, 5-bromodeoxyuridine (BrdU) and lysozyme (Lyso), showing that 72 h of  $10 \text{ mg kg}^{-1}$  rapamycin treatment causes a loss in RPS6 phosphorylation and BrdU positivity, and an increase in lysozyme staining in intestinal tumours (representative of 5 biological replicates). **f**, Representative H&E staining and IHC for BrdU showing that small, non-proliferative lesions remain after 30 days of  $10 \text{ mg kg}^{-1}$  rapamycin treatment (representative of 5 biological replicates). Scale bars,  $100 \mu\text{m}$ .

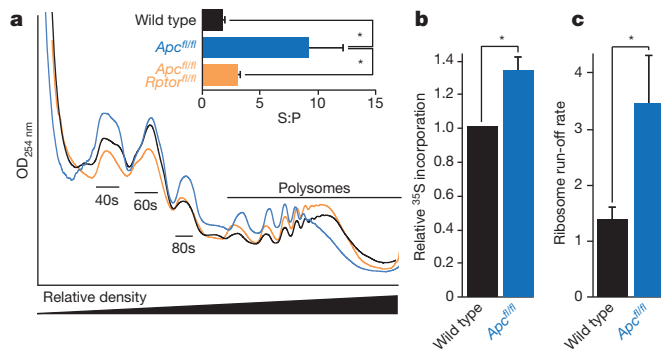
longer than controls (Fig. 2c, d and Extended Data Fig. 3c). We next analysed the tumours from these mice over a time course after rapamycin treatment. Treatment caused a loss of proliferation specifically within the tumours by 72 h, and an increase in the number of lysozyme-positive Paneth cells (Fig. 2e and Extended Data Fig. 3d, e). By 30 days, most tumours had shrunk considerably to small non-proliferative lesions that no longer contained Paneth cells (Fig. 2f and Extended Data Fig. 3f). Within the normal intestine there are two main cell populations that show high levels of Wnt signalling; the label-retaining/progenitor population and the Paneth cell population<sup>15</sup>. Our data suggest that treatment of mice with rapamycin causes the differentiation of the tumour's Wnt-high progenitor cells into the other Wnt-high fate in the intestine: namely non-proliferative Paneth-like cells. The cell-cycle arrest in these cells was examined by staining for p21, p16 and p53. No increase in these markers was observed, suggesting that a classical cell-cycle arrest pathway had not been engaged (Extended Data Fig. 4a). We reasoned that if mice were removed from rapamycin the tumours would regain proliferative capacity. Indeed, when rapamycin treatment was halted, signs of intestinal neoplasia were observed approximately 40–60 days later (Extended Data Fig. 3c). This suggested that intestinal adenoma stem cells were still present. Tumours from *Lgr5GFP-Cre<sup>ER</sup>* mice were stained to detect LGR5–GFP positivity. We found that, after rapamycin treatment, numerous LGR5-positive cells were still present, indicating that, although rapamycin treatment causes a regression of the lesions, the tumour-initiating cells remain (Extended Data Fig. 4b).

We next examined the mechanism of mTORC1 requirement after *Apc* loss. mTORC1 is known to regulate protein synthesis on multiple levels and most research has focused on two downstream effectors: 4EBP1 and S6K. A number of studies have suggested that translation initiation, via the 4EBP1–eIF4E axis, is the critical effector of mTOR in cancer<sup>6,16</sup>. However, it has been shown that rapamycin preferentially inhibits the phosphorylation of S6K over 4EBP1 (ref. 8), suggesting that 4EBP1-mediated inhibition of translation initiation may not be limiting in the context of *Apc* loss. To assess the changes in translational control in response to mTORC1 inhibition, we measured the polysomal distribution in wild-type, *Apc*-deficient and *Apc/Rptor*-deficient intestinal epithelial cells 4 days after gene deletion. *Apc* deletion resulted in a decrease in the number of polysomes, whereas *Apc/Rptor* co-deletion reversed this effect (Fig. 3a). The decrease in the number of polysomes after *Apc* deletion could suggest either reduced translation initiation (and, consequently, a lower overall level of translation) and/or a faster rate of translational elongation. Global translation rates were measured using an *in vitro* intestinal crypt culture model<sup>17</sup>. The *Apc*-deficient cells were shown to have increased <sup>35</sup>S-labelled methionine/cysteine incorporation compared with wild type, showing higher overall levels of protein synthesis (Fig. 3b). Unfortunately, *Rptor* deletion prevented the growth of crypts *in vitro* so this could not be assayed (Extended Data Fig. 5).

To measure the rate of translational elongation, an *in vitro* harringtonine run-off assay was performed<sup>18</sup>, as described in Methods. There was a >2.5-fold increase in ribosome run-off in crypts with *Apc* deletion compared with wild type (Fig. 3c and Extended Data Fig. 6a–d). This suggests that, after Wnt activation, elongation, rather than initiation, is rate limiting for protein synthesis and that mTORC1 must be activated to overcome this.

Cycloheximide (an inhibitor of elongation<sup>19</sup>) reduced proliferation associated with *Apc* deletion to a similar level to rapamycin (Extended Data Fig. 6e, f). While cycloheximide is acknowledged to inhibit elongation<sup>19</sup>, it must be emphasized that 72 h treatment could result in broad alterations in protein synthesis. However, the *Apc<sup>fl/fl</sup>*-specific loss of proliferation observed here provides ‘proof of principle’ to demonstrate that the modulation of protein synthesis may be useful as a chemotherapeutic strategy.

As most previous work has suggested that translation initiation downstream of 4EBP1 is limiting to cancer<sup>20</sup>, it was important to probe known effectors of mTORC1 in this system. Given the alteration of elongation rates, eEF2K, a known target of S6K<sup>21,22</sup> was of particular interest. eEF2K



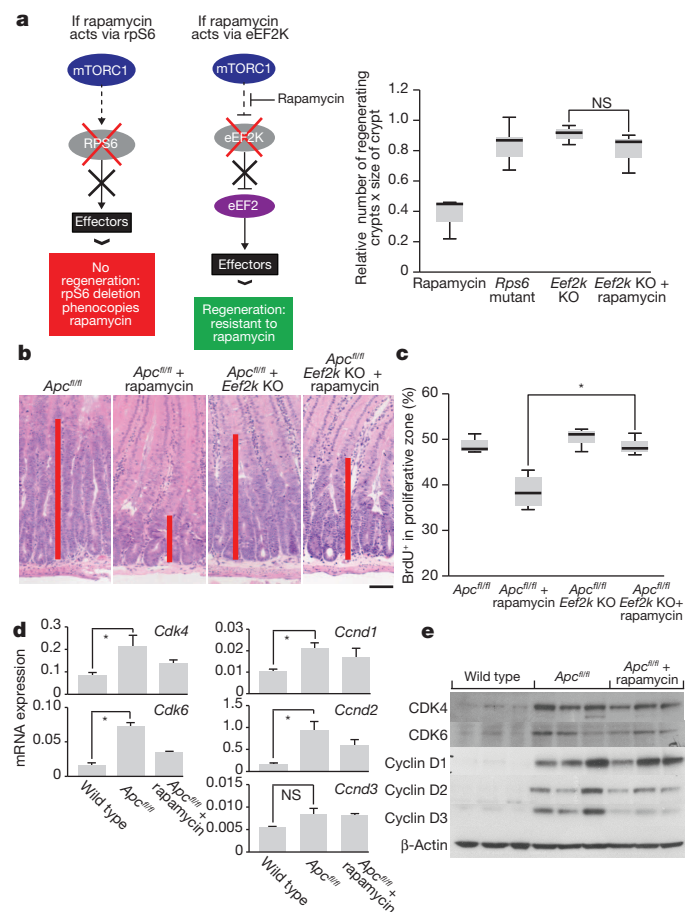
**Figure 3 | mTORC1 drives increased translational elongation.**

**a**, Representative polysome profiles of intestinal epithelial cells showing altered RNA distribution 96 h after *Apc* deletion. Bar graph represents the ratio of sub-polysomes compared with polysomes (S:P).  $OD_{254\text{ nm}}$ , optical density at 254 nm. Data are average  $\pm$  standard error of the mean (s.e.m.) ( $n = 3$  per group). \* $P$  value  $\leq 0.05$ , Mann–Whitney  $U$  test. **b**, Intestinal crypt culture was pulsed for 30 min with  $^{35}\text{S}$ -labelled methionine/cysteine. Incorporation of  $^{35}\text{S}$  into protein was quantified by scintillation counting and normalized to total protein. *Apc* deletion increases  $^{35}\text{S}$  incorporation. Data are average  $\pm$  s.e.m. ( $n = 3$  biological replicates per group). \* $P$  value  $\leq 0.05$ , Mann–Whitney  $U$  test. **c**, The ribosome run-off rate was measured by addition of the initiation inhibitor harringtonine to *ex vivo* crypts from wild-type and *Apc*-deleted mice. Harringtonine was added for 0 or 180 s and the increase in sub-polysomes relative to polysomes was calculated. This run-off rate represents the shift in S:P between the two time points, which is proportional to elongation speed. Data are average  $\pm$  s.e.m. ( $n = 3$  biological replicates per group). \* $P$  value  $\leq 0.05$ , Mann–Whitney  $U$  test. Also see Extended Data Fig. 6.

is a negative regulator of eEF2, giving mTORC1 the ability to promote translational elongation via S6K (ref. 23). Using multiple mouse knockout and knock-in alleles, we further dissected the downstream effectors of mTORC1 in intestinal regeneration. *S6k1/2* (also known as *Rps6kb1* and *Rps6kb2*, respectively) knockout decreased intestinal regeneration, while knockout of *Eif4ebp1/2* (coding 4EBP1 and 4EBP2, respectively) had no effect. As the 4EBP proteins are negative regulators of eIF4E, an increase (rather than a decrease) in regeneration may have been predicted, but this was not found. Moreover these intestines were still sensitive to rapamycin, demonstrating that rapamycin was acting via the mTORC1–S6K branch rather than the 4EBP1–eIF4E branch (Extended Data Fig. 7). We then used an *Eef2k*-null mouse and showed that after irradiation and treatment with rapamycin, these mice were now resistant to mTORC1 inhibition, confirming the importance of translational elongation (Fig. 4a). To ensure that S6K was not also acting through its more established effector, rpS6, we used an *Rps6* phospho-mutant that cannot be phosphorylated by S6K. This was unable to phenocopy *Rptor* deletion (Fig. 4a), showing that Wnt-driven regeneration requires increased translational elongation, mediated through mTORC1.

To prove that inhibition of eEF2K by S6K is required to allow increased eEF2 activity after *Apc* loss, we intercrossed *VillinCre<sup>ER</sup> Apc<sup>fl/fl</sup>* to *Eef2k<sup>-/-</sup>* mice and treated these with rapamycin. In contrast to *VillinCre<sup>ER</sup> Apc<sup>fl/fl</sup>* mice, these intestines were now resistant to the growth inhibitory effects of rapamycin (Fig. 4b, c). Tellingly, these mice no longer show an increase in the inhibitory phosphorylation of eEF2 after rapamycin treatment (Extended Data Fig. 8).

To assess whether the increased elongation after *Apc* deletion had differential effects on cell-cycle-regulating proteins, RNA and protein levels of several key cell-cycle regulators were tested (Fig. 4d, e). This analysis revealed that while there were increased RNA and protein levels of cyclin D1, cyclin D2, CDK4 and CDK6, cyclin D3 had increased protein levels in the absence of increased messenger RNA levels. Cyclin D3 protein levels were sensitive to rapamycin exposure, and this sensitivity depends on eEF2K (Extended Data Fig. 9). Additionally, ribosomes were shown to elongate approximately four times faster on cyclin D3 messages in *Apc*-deficient cells than in wild-type cells. No differences were detected in



**Figure 4 | mTORC1 signalling via eEF2K controls intestinal proliferation after Wnt signalling.**

**a**, Graphical representation of findings and boxplots showing that murine intestinal regeneration after irradiation implicates the mTORC1–S6K–eEF2K axis in Wnt-driven proliferation. Mice were exposed to 14 Gy  $\gamma$ -irradiation, and intestinal regeneration was calculated 72 h after exposure by examining the number and size of regenerating crypts relative to wild-type regenerating intestines. KO, knockout. Whiskers show maximum and minimum, black line shows median ( $n = 6$  per group). \* $P$  value  $\leq 0.05$ , Mann–Whitney  $U$  test. NS, not significant. **b**, **c**, Representative H&E and boxplot showing that *Eef2k* deletion confers resistance to  $10\text{ mg kg}^{-1}$  rapamycin treatment, 96 h after *Apc* deletion. Treatment began 24 h after induction. Red bar is graphical representation of crypt size. Whiskers show maximum and minimum, black line shows median ( $n = 3$  biological replicates per group). \* $P$  value  $< 0.05$ , Mann–Whitney  $U$  test. Scale bar,  $100\text{ }\mu\text{m}$ . **d**, Polymerase chain reaction with quantitative reverse transcription (qRT–PCR) of intestinal epithelial cells using primers for *Cdk4*, *Cdk6*, *Ccn1*, *Ccn2* and *Ccn3*. *Ccn3* is not regulated at the transcriptional level. Data were normalized to *Gapdh*. Data are average  $\pm$  s.e.m. ( $n = 3$  biological replicates per group). \* $P$  value  $\leq 0.05$ , Mann–Whitney  $U$  test. NS, not significant. **e**, Western blot analysis of intestinal epithelial cells from each group. Antibodies to CDK4, CDK6, cyclin D1, cyclin D2, cyclin D3 and  $\beta$ -actin are shown. Each well represents a different mouse from the relevant group, and the samples are the same as those used for the qRT–PCR.

other messages tested (Extended Data Fig. 10). Taken together, these data suggest that cyclin D3 is translationally regulated at the level of elongation, consistent with previous reports<sup>24,25</sup>. The contribution of cyclin D3 to the proliferative phenotype remains to be elucidated.

We report that mTORC1 is an essential downstream effector of Wnt signalling in the intestine. We show that intestinal proliferation associated with Wnt signalling requires the mTORC1–S6K–eEF2K–eEF2 axis, and that the resulting increase in the rate of elongation of specific polypeptides overcomes a limiting translational step. Our work highlights key functional roles for eEF2K and translational elongation in the control of the initiation of cancer and adenomatous proliferation. The

importance of elongation in this context has been suggested in a small number of publications, but this study provides key *in vivo* evidence<sup>26–28</sup>. Finally, we have also shown that targeting mTOR and translational control may be a viable strategy for chemoprevention of colorectal carcinoma in high-risk patients, and treatment of early stage disease. Indeed, recent studies have suggested that the chemopreventative agents aspirin and mesalazine also target mTOR<sup>29,30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 4 October 2013; accepted 26 September 2014.**

**Published online 5 November 2014.**

- Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87**, 159–170 (1996).
- Korinek, V. *et al.* Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC<sup>−/−</sup> colon carcinoma. *Science* **275**, 1784–1787 (1997).
- Ashton, G. H. *et al.* Focal adhesion kinase is required for intestinal regeneration and tumorigenesis downstream of Wnt/c-Myc signaling. *Dev. Cell* **19**, 259–269 (2010).
- Fujishita, T., Aoki, K., Lane, H. A., Aoki, M. & Taketo, M. M. Inhibition of the mTORC1 pathway suppresses intestinal polyp formation and reduces mortality in Apc<sup>Δ716</sup> mice. *Proc. Natl Acad. Sci. USA* **105**, 13544–13549 (2008).
- Gulhati, P. *et al.* Targeted inhibition of mammalian target of rapamycin signaling inhibits tumorigenesis of colorectal cancer. *Clin. Cancer Res.* **15**, 7207–7216 (2009).
- Pourdehnad, M. *et al.* Myc and mTOR converge on a common node in protein synthesis control that confers synthetic lethality in Myc-driven cancers. *Proc. Natl Acad. Sci. USA* **110**, 11988–11993 (2013).
- Martineau, Y. *et al.* Pancreatic tumours escape from translational control through 4E-BP1 loss. *Oncogene* **33**, 1367–1374 (2014).
- Jiang, Y. P., Ballou, L. M. & Lin, R. Z. Rapamycin-insensitive regulation of 4E-BP1 in regenerating rat liver. *J. Biol. Chem.* **276**, 10943–10951 (2001).
- Bach, S. P., Renahan, A. G. & Potten, C. S. Stem cells: the intestinal stem cell as a paradigm. *Carcinogenesis* **21**, 469–476 (2000).
- Bernal, N. P. *et al.* Evidence for active Wnt signaling during postresection intestinal adaptation. *J. Pediatr. Surg.* **40**, 1025–1029 (2005).
- Ireland, H. *et al.* Inducible Cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of  $\beta$ -catenin. *Gastroenterology* **126**, 1236–1246 (2004).
- Muncan, V. *et al.* Rapid loss of intestinal crypts upon conditional deletion of the Wnt/Tcf-4 target gene *c-Myc*. *Mol. Cell. Biol.* **26**, 8418–8426 (2006).
- Zoncu, R., Efeyan, A. & Sabatini, D. M. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nature Rev. Mol. Cell Biol.* **12**, 21–35 (2011).
- Yilmaz, Ö. H. *et al.* mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* **486**, 490–495 (2012).
- Farin, H. F., Van Es, J. H. & Clevers, H. Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells. *Gastroenterology* **143**, 1518–1529 (2012).
- She, Q. B. *et al.* 4E-BP1 is a key effector of the oncogenic activation of the AKT and ERK signaling pathways that integrates their function in tumors. *Cancer Cell* **18**, 39–51 (2010).
- Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
- Fresno, M., Jimenez, A. & Vazquez, D. Inhibition of translation in eukaryotic systems by harringtonine. *Eur. J. Biochem.* **72**, 323–330 (1977).
- Schneider-Poetsch, T. *et al.* Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nature Chem. Biol.* **6**, 209–217 (2010).
- Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485**, 55–61 (2012).
- Richter, J. D. & Sonenberg, N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433**, 477–480 (2005).
- Browne, G. J. & Proud, C. G. A novel mTOR-regulated phosphorylation site in elongation factor 2 kinase modulates the activity of the kinase and its binding to calmodulin. *Mol. Cell. Biol.* **24**, 2986–2997 (2004).
- Ryazanov, A. G., Shestakova, E. A. & Natapov, P. G. Phosphorylation of elongation factor 2 by EF-2 kinase affects rate of translation. *Nature* **334**, 170–173 (1988).
- Gorshtein, A. *et al.* Mammalian target of rapamycin inhibitors rapamycin and RAD001 (everolimus) induce anti-proliferative effects in GH-secreting pituitary tumor cells *in vitro*. *Endocr. Relat. Cancer* **16**, 1017–1027 (2009).
- Gutzkow, K. B. *et al.* Cyclic AMP inhibits translation of cyclin D3 in T lymphocytes at the level of elongation by inducing eEF2-phosphorylation. *Cell. Signal.* **15**, 871–881 (2003).
- Firczuk, H. *et al.* An *in vivo* control map for the eukaryotic mRNA translation machinery. *Mol. Syst. Biol.* **9**, 635 (2013).
- Hussey, G. S. *et al.* Identification of an mRNP complex regulating tumorigenesis at the translational elongation step. *Mol. Cell* **41**, 419–431 (2011).
- Nakamura, J. *et al.* Overexpression of eukaryotic elongation factor eEF2 in gastrointestinal cancers and its involvement in G2/M progression in the cell cycle. *Int. J. Oncol.* **34**, 1181–1189 (2009).
- Din, F. V. *et al.* Aspirin inhibits mTOR signaling, activates AMP-activated protein kinase, and induces autophagy in colorectal cancer cells. *Gastroenterology* **142**, 1504–1515 (2012).
- Baan, B. *et al.* 5-Aminosalicylic acid inhibits cell cycle progression in a phospholipase D dependent manner in colorectal cancer. *Gut* **61**, 1708–1715 (2012).

**Acknowledgements** W.J.F. is funded by AICR. O.J.S. is funded by Cancer Research UK, European Research Council Investigator Grant (COLONCAN) and the European Union Seventh Framework Programme FP7/2007–2013 under grant agreement number 278568. M.B. is a Medical Research Council Senior Fellow. The authors acknowledge P. Cammareri, J. Morton and C. Murgia for proofreading of the manuscript.

**Author Contributions** O.J.S., A.E.W. and W.J.F. designed the project. W.J.F., R.A.R., T.J. and S.R. performed breeding and phenotypic analysis of mice; W.J.F., T.J.J. and J.R.P.K. performed translational analysis; M.N.H., A.G.R., N.S., O.M., A.S., J.B.C., M.V., D.J.H., K.B.M., S.A.K., K.M.D., C.J., H.A.C. and M.P. provided advice and material; W.J.F., O.J.S., A.E.W. and M.B. wrote and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to O.J.S. ([o.sansom@beatson.gla.ac.uk](mailto:o.sansom@beatson.gla.ac.uk)).



## METHODS

**Mouse colonies.** All experiments were performed according with UK Home Office regulations (licence 60/4183) which undergoes local ethical review at Glasgow University. Outbred male mice from 6 to 12 weeks of age were used. The majority of the work was performed on C57BL/6 mice: *AhCre*, *Apc<sup>fl/fl</sup>*, *Apc<sup>Min/+</sup>*, *Rptor<sup>fl/fl</sup>*, *S6k1/2* knockout and *Rps6<sup>mut</sup>* mice were all C57BL/6J. Some treatment experiments were performed on mice that were only three generations C57BL/6 (*Lgr5Cre<sup>ER</sup> Apc<sup>fl/fl</sup>*).

The alleles used were as follows: *VillinCre<sup>ER</sup>* (ref. 31), *AhCre* (ref. 11), *Lgr5Cre<sup>ER</sup>* (ref. 32), *Apc<sup>S80</sup>* (ref. 33), *Apc<sup>Min/+</sup>* (ref. 34), *Myd<sup>fl/fl</sup>* (ref. 35), *Rptor<sup>fl/fl</sup>* (ref. 36), *ROSA-tdRFP* (ref. 37), *Eif4ebp1* knockout (ref. 38), *Eif4ebp2* knockout (ref. 39), *S6k1* knockout (ref. 40), *S6k2* knockout (ref. 41), *Eef2k* knockout (ref. 42), *Rps6<sup>mut</sup>* (ref. 43). Recombination by *VillinCre<sup>ER</sup>* was induced with one intraperitoneal (i.p.) injection of 80 mg kg<sup>-1</sup> tamoxifen on day 0 and day 1. Analysis of *VillinCre<sup>ER</sup>*-induced mice was at day 4 after induction. Red fluorescent protein (RFP) analysis was carried out by inducing recombination by *AhCre* using a single i.p. injection of 80 mg kg<sup>-1</sup>  $\beta$ -naphthoflavone. RFP visualization was carried out on day 4. Mice carrying the *Lgr5Cre<sup>ER</sup>* transgene were given one i.p. injection of 120 mg kg<sup>-1</sup> tamoxifen.

For regeneration experiments, mice were exposed to  $\gamma$ -irradiation from caesium-137 sources. This delivered  $\gamma$ -irradiation at 0.423 Gy min<sup>-1</sup>.

Rapamycin treatment was performed using a daily i.p. injection of 10 mg kg<sup>-1</sup> (refs 43, 44) in 5% Tween80 and 5% polyethylene glycol in PBS. Cycloheximide treatment was performed using a daily i.p. injection of 35 mg kg<sup>-1</sup> in PBS.

In accordance with the 3Rs, the smallest sample size was chosen that could give a significant difference. Given the robust phenotype of the *Apc<sup>fl/fl</sup>*, and our prediction that mTOR was essential, the minimum sample size assuming no overlap in control versus experimental is three animals.

No randomization was used and the experimenter was blinded to drugs and genotypes.

**IHC.** Standard IHC techniques were used throughout this study. Antibody concentrations used were as follows: phospho-rpS6<sup>Ser235/236</sup> (1:800; Cell Signaling 4858), phospho-4EBP1<sup>Thr37/46</sup> (1:500; Cell Signaling 2855), phospho-eEF2<sup>Thr56</sup> (1:500; Novus Biologicals NB100-92518), c-MYC (1:200; Santa Cruz sc-764),  $\beta$ -catenin (1:50; BD Biosciences 610154), BrdU (1:200; BD Biosciences 347580), lysozyme (1:150; Dako A099), GFP (1:1,000; Abcam ab6556), p21 (1/4; CNIO Madrid), p16 (1:400; Santa Cruz sc1661), p53 (1/150; Vector Laboratories VPP956). For each antibody, staining was performed on at least three mice of each genotype. Representative images are shown for each staining.

**Assaying apoptosis, mitosis and proliferation *in vivo*.** Apoptosis and mitotic index were scored from H&E-stained sections as previously described<sup>11</sup>. Proliferation levels were assessed by measuring BrdU incorporation. Mice were injected with 250  $\mu$ l of BrdU (Amersham Biosciences) 2 h before being killed. Immunohistochemical staining for BrdU was then performed using an anti-BrdU antibody. For each analysis, 25 full crypts were scored from at least three mice of each genotype.

**Intestinal epithelium extraction.** To generate tissue for polysomal profile analysis, 10-cm portions of intestine were flushed with 0.1 mg ml<sup>-1</sup> cycloheximide (Sigma) in PBS and inverted over a glass rod to expose the epithelial surface. Intestines were incubated in 0.1 mg ml<sup>-1</sup> cycloheximide in HBSS (Gibco) with 10 mM EDTA for 5 min at 37 °C followed by 5 min of vigorous shaking. Intestines were transferred to 0.1 mg ml<sup>-1</sup> cycloheximide in PBS and incubated for a further 5 min at 4 °C, followed by 5 min of vigorous shaking. This fraction contained intestinal crypts and was used for downstream analysis.

**Sucrose density ultracentrifugation.** Intestinal epithelial cells were lysed in ice-cold 300 mM NaCl, 15 mM MgCl<sub>2</sub>, 15 mM Tris (pH 7.5) containing 500 units ml<sup>-1</sup> RNasin, 1 mg ml<sup>-1</sup> heparin sulphate and 0.1 mg ml<sup>-1</sup> cycloheximide supplemented with 0.1% (v/v) Triton X-100. Post-nuclear lysates were layered on 10 ml 10–50% (w/v) sucrose gradients of the same buffer omitting Triton X-100. Gradients were centrifuged at 38,000 r.p.m. for 2 h at 4 °C in a SW40Ti rotor (Beckman Coulter) and separated through a live OD<sub>254 nm</sub> ultraviolet spectrometer (Isco). Comparison of peak abundance was based on the area under the curve.

**Crypt culture.** Mouse small intestines were opened longitudinally and washed with PBS. Crypts were isolated as previously described<sup>17</sup>. Isolated crypts were mixed with 50  $\mu$ l of Matrigel (BD Bioscience), plated in 24-well plates in Advanced DMEM/F12 with Noggin (100 ng ml<sup>-1</sup>, Peprotech). Wild-type crypts were also supplemented with R-spondin (500 ng ml<sup>-1</sup>, R&D Systems). Growth factors were added every other day. Sphere formation was scored 7 days after plating, by counting the number of spheres present per well.

**Determination of protein synthesis rates.** Cells were treated with 30  $\mu$ Ci ml<sup>-1</sup> <sup>35</sup>S-methionine label (Hartmann Analytic) for 30 min then harvested and lysed. Protein was precipitated onto filter paper (Whatmann) by addition of trichloroacetic acid to 12.5% and washed with 70% ethanol then acetone. Scintillation was read from dried filter paper in triplicate for each experimental condition (National Diagnostics). Total protein content was determined by bicinchoninic acid (BCA) assay (Pierce) for standardization between conditions.

**Harringtonine run-off assay.** Harringtonine inhibits *de novo* translational initiation, allowing ribosomes engaged in elongation to run-off their messages while limiting re-initiation post-termination. Harringtonine was added for 0 or 180 s and the increase in sub-polysomes relative to polysomes was calculated. This run-off rate represents the shift in S:P between the two time points, which is proportional to elongation speed. Crypt cultures were treated with 2  $\mu$ g ml<sup>-1</sup> Harringtonine (Insight Biotechnology) and at set time periods (0 and 180 s) 0.1 mg ml<sup>-1</sup> cycloheximide was added. Cells were scraped into PBS at 4 °C and prepared for sucrose gradient ultracentrifugation as previously described.

**Western blotting.** Snap-frozen intestinal epithelial tissue (50–100 mg) was homogenized using the Precellys 24 (Stretton Scientific) in 500  $\mu$ l of Ripa-lysis buffer. Protein concentrations were determined using a BCA Protein Assay Kit (Pierce). Equal amounts of cellular protein (30  $\mu$ g) were separated on a 4–12% gradient gel (Novex) and subsequently transferred to a PVDF membrane (Amersham). Total protein was visualized with Poinceau (Sigma). After blocking the membranes in TBS containing 5% BSA (Sigma), 0.02% Triton X-100 for 1 h, primary antibodies were added in block solution at the following dilutions: CDK4 (Santa Cruz SC-260, 1:1,000), CDK6 (Cell Signaling 3136, 1:1,000), cyclin D1 (Cell Signaling 2926, 1:2,000), cyclin D2 (Cell Signaling 3741, 1:1,000), cyclin D3 (Cell Signaling 2936, 1:2,000), eEF2K (Cell Signaling 3692, 1:1,000) and  $\beta$ -actin (Sigma A2228, 1:5,000). After washing, the appropriate HRP-conjugated secondary goat antibodies (Dako) were added diluted 1:10,000 in block for 1 h. Antibody binding was detected using ECL Western Blotting Substrate (Pierce). Primary antibody incubations were carried out at 4 °C overnight. Remaining incubations were carried out at room temperature.

**RNA isolation.** Snap-frozen intestinal epithelial tissue was homogenized and RNA was extracted using the TRIzol method (Ambion).

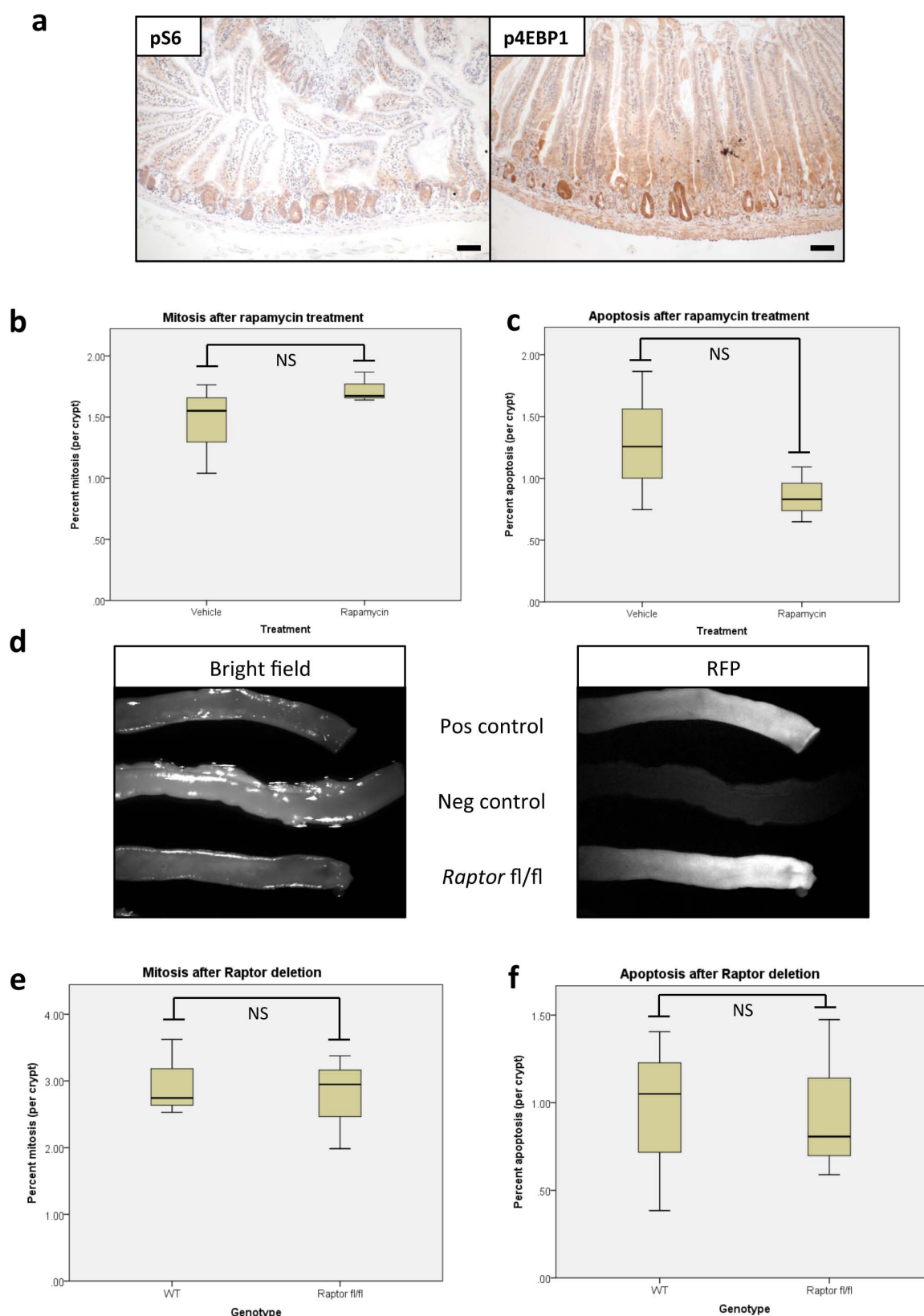
**qPCR.** One microgram of RNA was reverse transcribed to cDNA using a Quantitect Reverse Transcription Kit (Qiagen) in a reaction volume of 20  $\mu$ l. qPCR was performed on each sample in triplicate in a 20  $\mu$ l reaction mixture containing 10  $\mu$ l of 2 $\times$  DyNAmo HS master mix (Thermo Scientific), 0.5  $\mu$ M of each of the primers (detailed later) and 2  $\mu$ l cDNA generated previously. The reaction mixture without a template was run in triplicate as a control. The reaction conditions were as follows: 95 °C for 15 min, followed by 40 cycles of three steps consisting of denaturation at 94 °C for 15 s, primer annealing at 60 °C for 30 s, and primer extension at 72 °C for 30 s. A melting curve analysis was performed from 70 °C to 95 °C in 0.3 °C intervals. *Gapdh* was used to normalize for differences in RNA input.

**qRT-PCR primers.** qRT-PCR primers were as follows. *Ccnd1* forward, 5'-GAGA AGTTGTGATCTACACTG-3'; *Ccnd1* reverse, 5'-AAATGAACCTTCACATCT GTGGC-3'; *Ccnd2* forward, 5'-CTACCGACTTCAAGTTTGCC-3'; *Ccnd2* reverse, 5'-GCTTTGAGACAATCCACATCAG-3'; *Cdk4* forward, 5'-AATGTTGTACG GCTGATGGA-3'; *Cdk4* reverse, 5'-AGAACTGACGCATTAGATCCT-3'; *Cdk6* forward, 5'-GGCGTACCCACAGAAACCATA-3'; *Cdk6* reverse, 5'-AGGTAAG GGCCATCTGAAAACCT-3'; *Ccnd3* forward, 5'-CGAGCCTCTACTTCCAGT G-3'; *Ccnd3* reverse, 5'-GGACAGGTAGCGATCCAGGT-3'; *Rps6* forward, 5'-A GCTCCGCACCTTCTATGAGA-3'; *Rps6* reverse, 5'-GGGAAAACCTTGCTTG TCATTC-3'; *Rps21* forward, 5'-GTCCATCCAGATGAACGTGG-3'; *Rps21* reverse, 5'-CCATCAGCCTTAGCCAATCGG-3'.

- El Marjoui, F. *et al.* Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* **39**, 186–193 (2004).
- Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
- Shibata, H. *et al.* Rapid colorectal adenoma formation initiated by conditional targeting of the *Apc* gene. *Science* **278**, 120–123 (1997).
- Moser, A. R., Pitot, H. C. & Dove, W. F. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**, 322–324 (1990).
- de Alboran, I. M. *et al.* Analysis of C-MYC function in normal cells via conditional gene-targeted mutation. *Immunity* **14**, 45–55 (2001).
- Polak, P. *et al.* Adipose-specific knockout of raptor results in lean mice with enhanced mitochondrial respiration. *Cell Metab.* **8**, 399–410 (2008).
- Luche, H., Weber, O., Nageswara Rao, T., Blum, C. & Fehling, H. J. Faithful activation of an extra-bright red fluorescent protein in “knock-in” Cre-reporter mice ideally suited for lineage tracing studies. *Eur. J. Immunol.* **37**, 43–53 (2007).
- Tsukiyama-Kohara, K. *et al.* Adipose tissue reduction in mice lacking the translational inhibitor 4E-BP1. *Nature Med.* **7**, 1128–1132 (2001).
- Banko, J. L. *et al.* The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. *J. Neurosci.* **25**, 9581–9590 (2005).
- Shima, H. *et al.* Disruption of the *p70<sup>s6k</sup>/p85<sup>s6k</sup>* gene reveals a small mouse phenotype and a new functional S6 kinase. *EMBO J.* **17**, 6649–6659 (1998).
- Ryazanov, A. G. Elongation factor-2 kinase and its newly discovered relatives. *FEBS Lett.* **514**, 26–29 (2002).
- Ruvinsky, I. *et al.* Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes Dev.* **19**, 2199–2211 (2005).

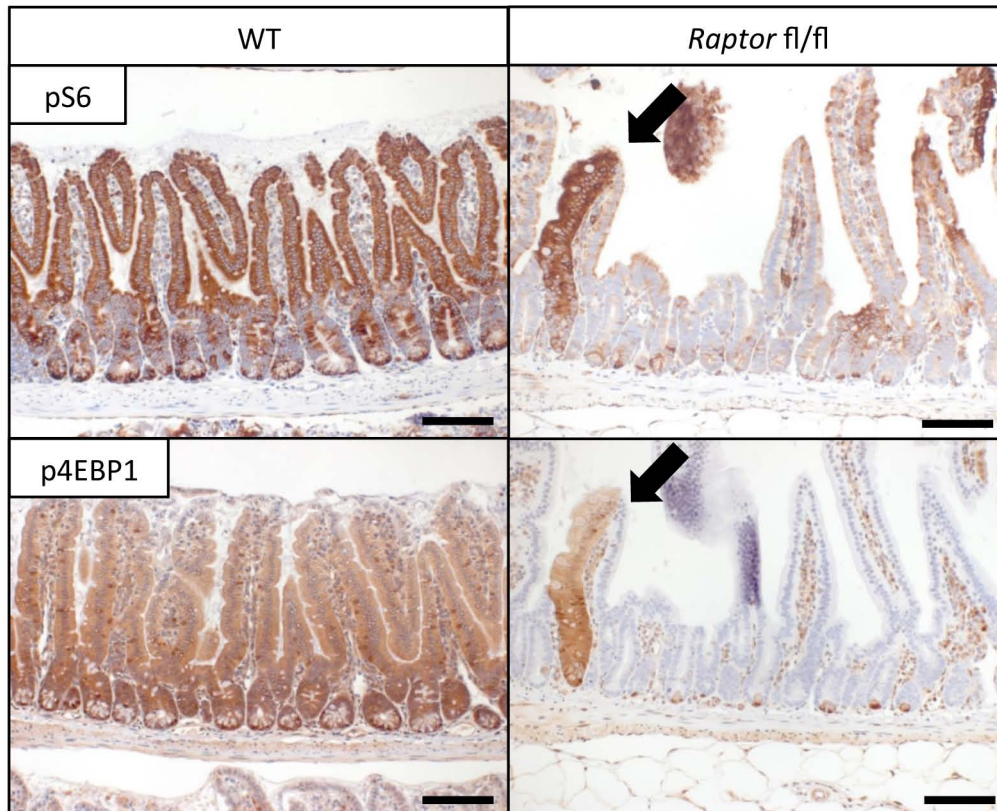
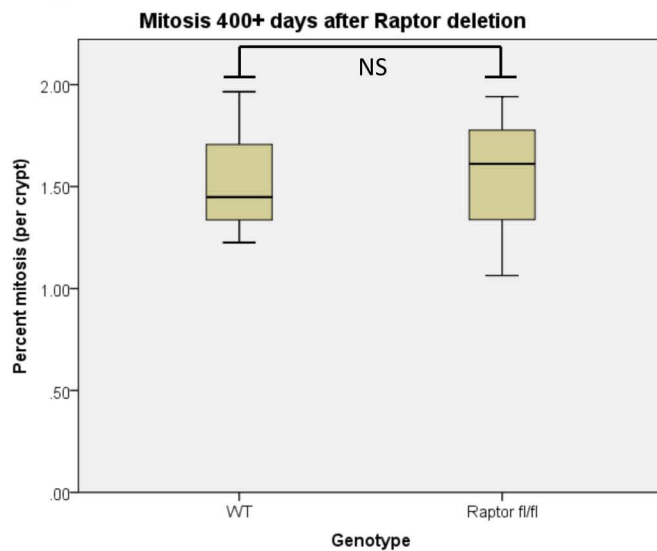
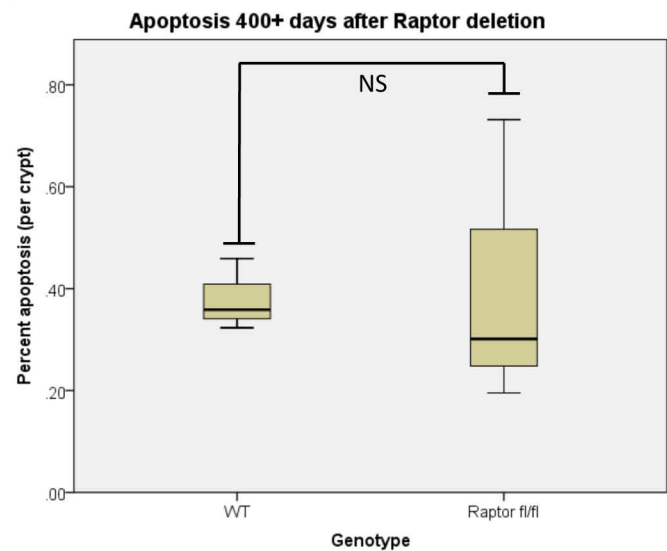


43. Sarbassov, D. D. *et al.* Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. *Mol. Cell* **22**, 159–168 (2006).
44. Sengupta, S., Peterson, T. R., Laplante, M., Oh, S. & Sabatini, D. M. mTORC1 controls fasting-induced ketogenesis and its modulation by ageing. *Nature* **468**, 1100–1104 (2010).



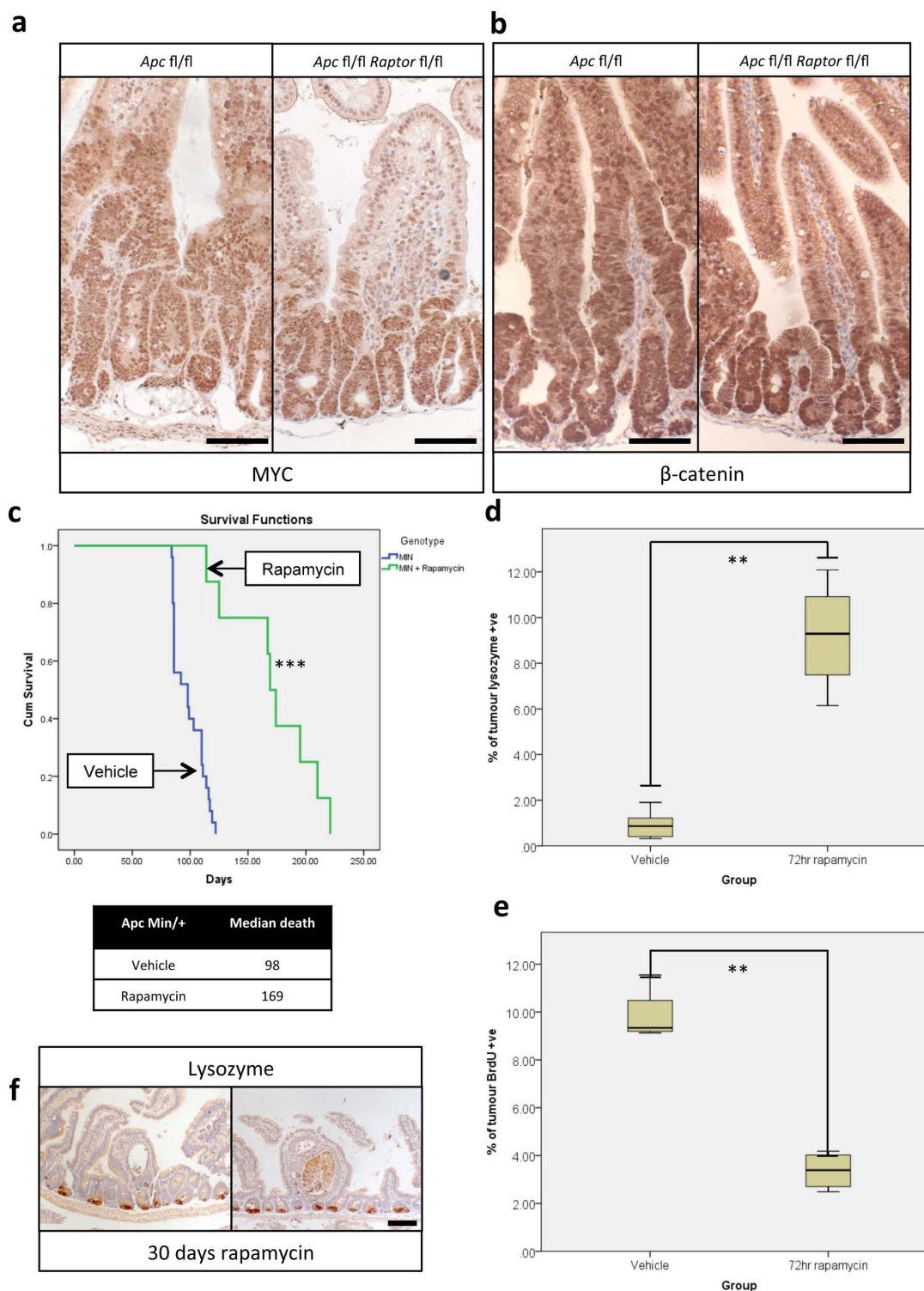
**Extended Data Figure 1 | mTORC1 is activated following Wnt-signal and its inhibition does not affect homeostasis.** **a**, Representative IHC of phospho-RPS6 (pS6) and phospho-4EBP1 (p4EBP1) show mTORC1 activity during intestinal regeneration, 72 h after 14 Gy  $\gamma$ -irradiation (representative of 5 biological replicates). **b**, **c**, Boxplots demonstrating that 72 h of 10 mg kg<sup>-1</sup> rapamycin treatment does not alter mitosis or apoptosis in normal intestinal crypts. Whiskers show maximum and minimum, black line shows median ( $n = 4$  per group). NS, not significant, Mann-Whitney  $U$  test. **d**, Intestines

imaged on OV100 microscope, 96 h after induction, for red fluorescent protein (RFP). Tissue without the ROSA-tdRFP reporter (Neg control) show no RFP positivity, while the positive control (Pos control) and *Raptor*-deleted intestines show high RFP positivity (representative of 3 biological replicates). **e**, **f**, Boxplot showing that *Raptor* deletion does not affect mitosis or apoptosis rates in intestinal crypts, 96 h after induction. Whiskers show maximum and minimum, black line shows median ( $n = 4$  per group). NS, not significant, Mann-Whitney  $U$  test. Scale bars, 100  $\mu$ m.

**a****b****c**

**Extended Data Figure 2 | *Raptor* deletion is maintained in the small intestine.** **a**, Representative IHC of phospho-RPS6 (pS6) and phospho-4EBP1 (p4EBP1) shows maintained loss of mTORC1 signalling 400+ days after *Raptor* deletion. Arrows indicate unrecaptured escaper crypts that still show active mTORC1 signalling (representative of 5 biological replicates). **b**, **c**, Boxplots

showing that mitosis and apoptosis are unchanged 400+ days after *Raptor* deletion. Mitosis and apoptosis were counted on H&E sections and are quantified as percent mitosis or apoptosis per crypt. Whiskers show maximum and minimum, black line shows median ( $n = 5$  per group). NS, not significant, Mann-Whitney  $U$  test. Scale bars, 100  $\mu$ m.

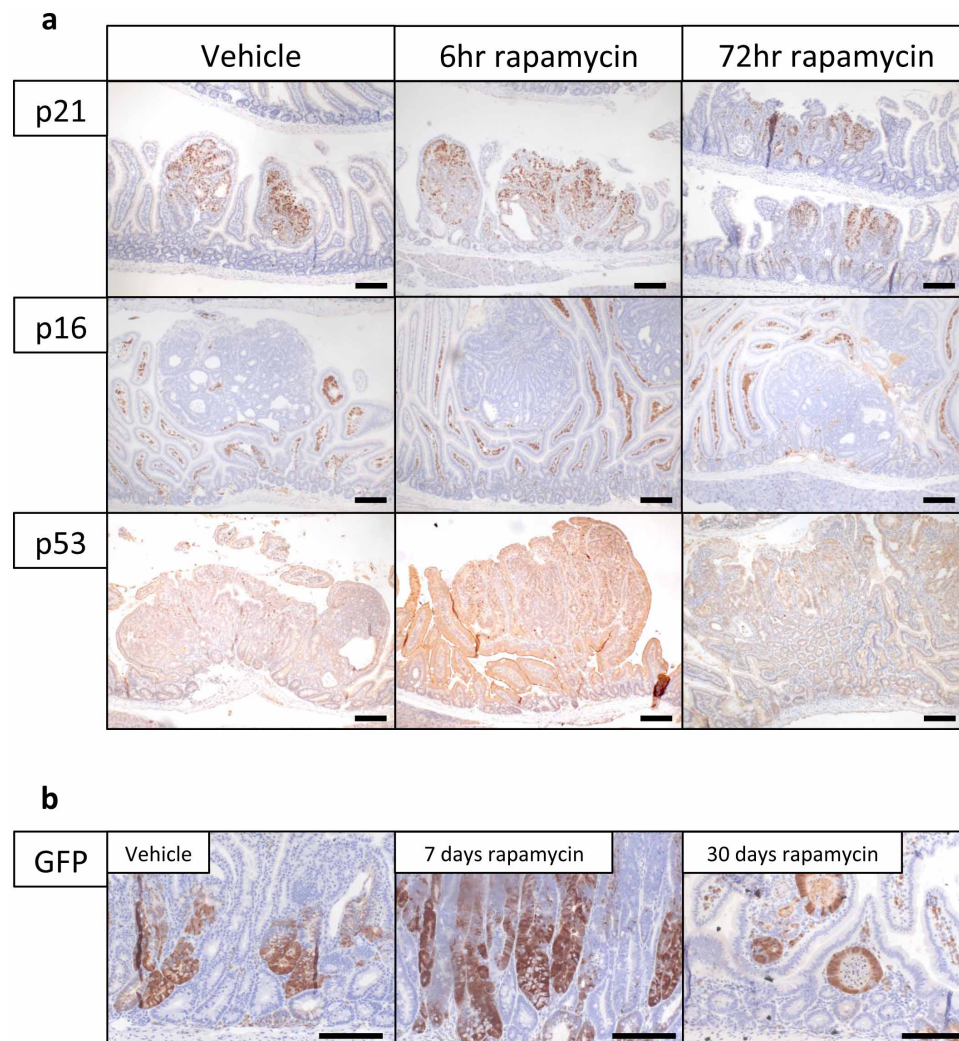


### Extended Data Figure 3 | Wnt signalling is still active after *Raptor* deletion and rapamycin treatment causes regression of established tumours.

**a, b**, Representative IHC of MYC and  $\beta$ -catenin showing high MYC levels and nuclear localization of  $\beta$ -catenin 96 h after *Apc* and *Apc/Raptor* deletion, demonstrating active Wnt signalling. Nuclear staining (as opposed to membranous staining) of  $\beta$ -catenin is indicative of active Wnt signalling. Scale bar, 100  $\mu$ m (representative of 3 biological replicates). **c**, Kaplan-Meier survival curve of *Apc*<sup>Min/+</sup> mice treated with rapamycin when showing signs of intestinal neoplasia. Rapamycin treatment (10 mg kg<sup>-1</sup>) started when mice showed signs of intestinal disease, and was withdrawn after 30 days. Animals continued to be observed until signs of intestinal neoplasia. Death of animals in the rapamycin group almost always occurred after rapamycin withdrawal ( $n = 8$  per group). \*\*\* $P$  value  $\leq 0.001$ , log-rank test. **d**, Boxplot showing that

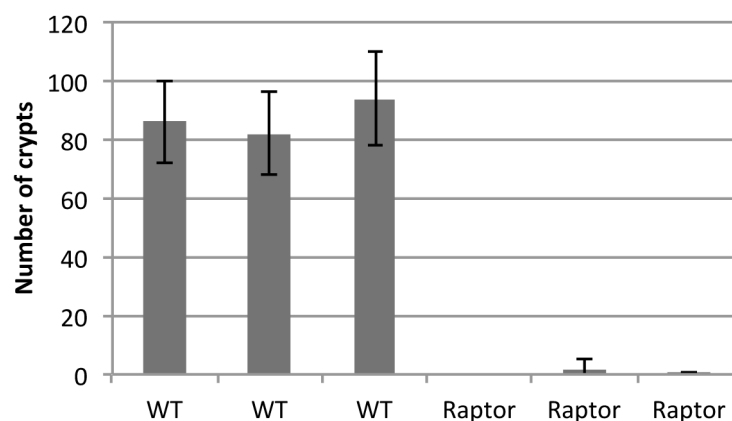
72 h 10 mg kg<sup>-1</sup> rapamycin treatment causes an increase in lysozyme-positive cells in tumours. Percentage lysozyme positivity within tumours was calculated using ImageJ software (<http://imagej.nih.gov/ij/>). Whiskers show maximum and minimum, black line shows median (10 tumours from each of 5 mice per group were measured). \*\* $P$  value  $\leq 0.014$ , Mann-Whitney  $U$  test. **e**, Boxplot showing that 72 h 10 mg kg<sup>-1</sup> rapamycin treatment causes a decrease in BrdU positivity within tumours. Percentage BrdU positivity within tumours was calculated using ImageJ software. Whiskers show maximum and minimum, black line shows median (10 tumours from each of 5 mice per group were measured). \*\*\* $P$  value  $\leq 0.021$ , Mann-Whitney  $U$  test. **f**, Representative IHC of lysozyme, showing a lack of lysozyme-positive paneth cells in remaining cystic tumours after 30 days of 10 mg kg<sup>-1</sup> rapamycin treatment. Scale bars, 100  $\mu$ m (representative of 5 biological replicates).





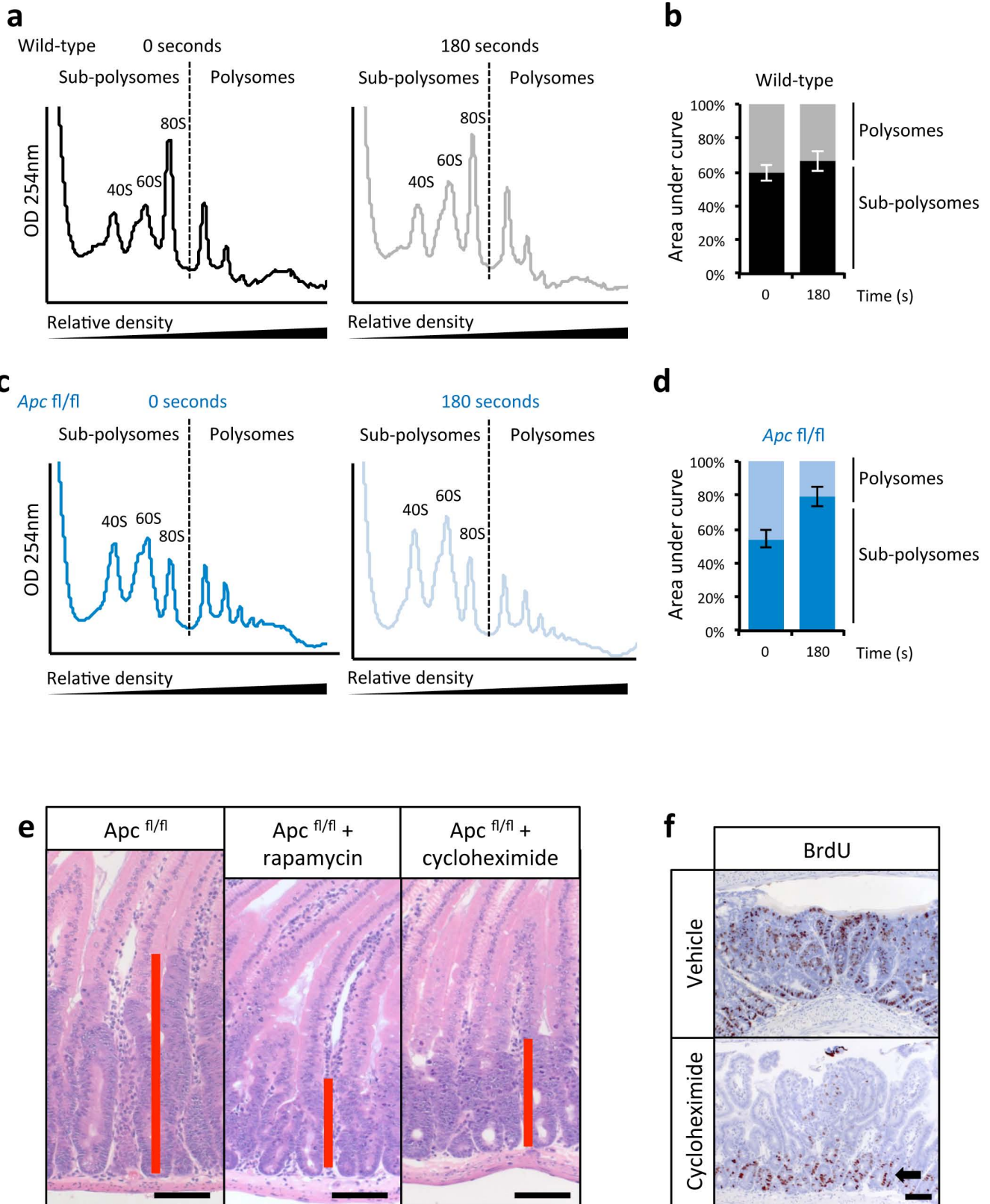
**Extended Data Figure 4 | IHC after rapamycin treatment.** **a**, Representative IHC of p21, p16 and p53 after 6 h and 72 h of  $10 \text{ mg kg}^{-1}$  rapamycin treatment. Staining shows no induction of these proteins in tumours after rapamycin treatment (representative of 5 biological replicates). **b**, Representative IHC for

LGR5–GFP showing high numbers of LGR5-positive cells after 7 and 30 days of  $10 \text{ mg kg}^{-1}$  daily rapamycin treatment (representative of 5 biological replicates). Scale bars,  $100 \mu\text{m}$ .

**a**

**Extended Data Figure 5 | *Rptor* deletion in the intestinal crypt is lethal *in vitro*.** **a**, Graph showing that *Rptor* deletion prevents intestinal crypts from growing *ex vivo*. Intestinal crypts were isolated and cultured as previously

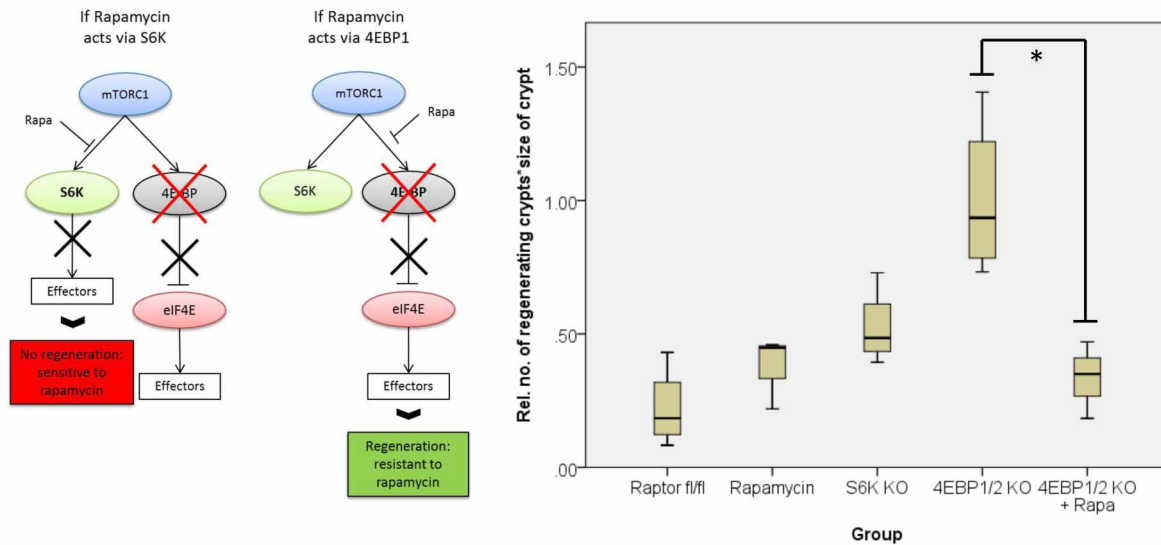
described<sup>17</sup>, 96 h after *Cre* induction. Number of viable organoids was counted by eye 72 h after crypt isolation. WT, wild type. Data are average  $\pm$  standard deviation ( $n = 3$  biological replicates per group).



**Extended Data Figure 6 | *Apc* deletion increases translational elongation rates and cycloheximide treatment phenocopies rapamycin treatment.**

**a**, Representative polysome profiles from wild-type *ex vivo* crypts incubated with harringtonine for 0 s (left) and 180 s (right) before harvest ( $n = 3$  per time point). **b**, The areas under the sub-polysome (40S, 60S and 80S) and polysome sections as indicated by the dashed lines in **a** were quantified and expressed as a percentage of their sum. Data in the bar graph are the average  $\pm$  s.e.m. ( $n = 3$  per time point). **c**, **d**, Data are shown for *Apc*-deleted

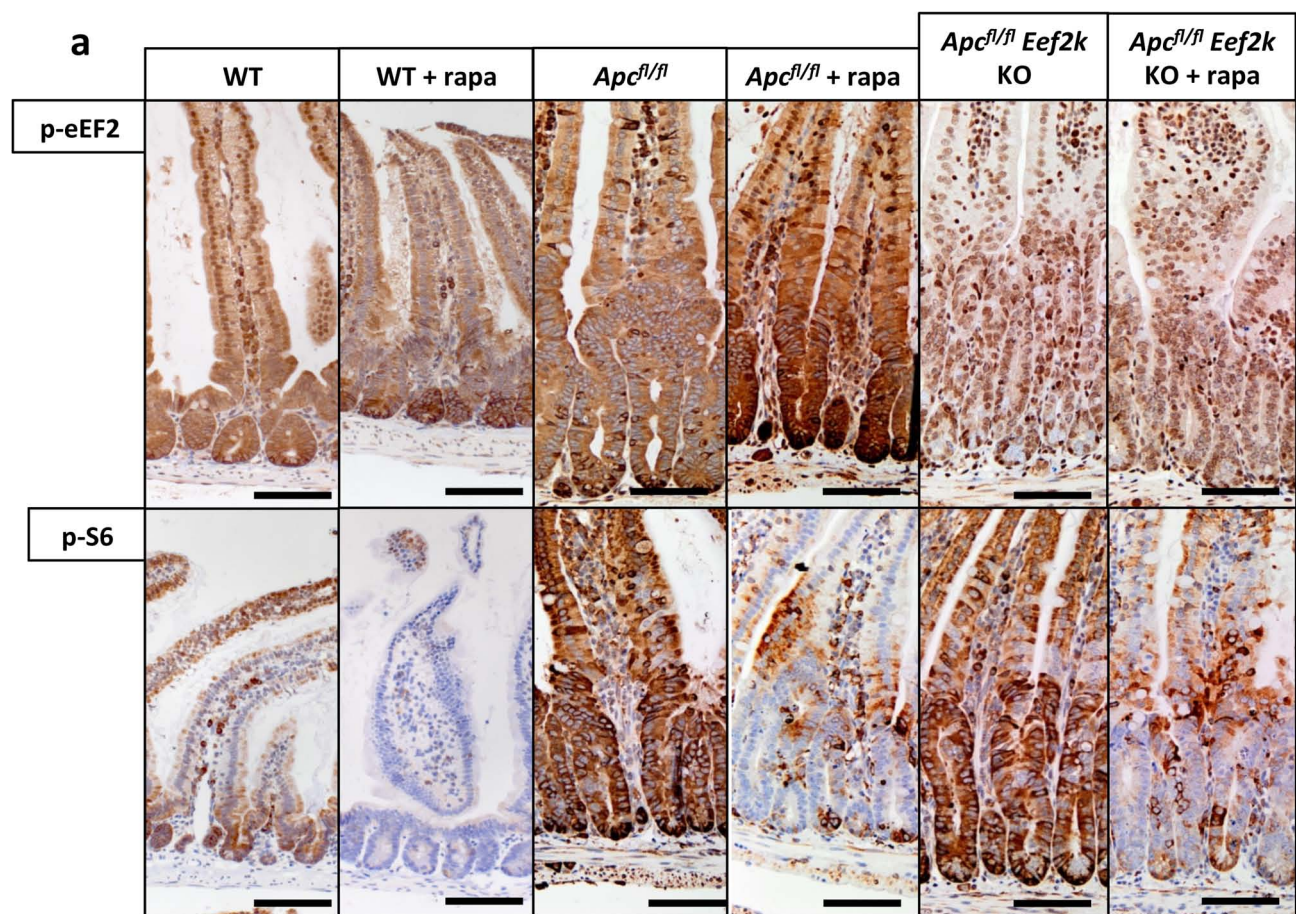
crypts, as for wild type in **b** and **c** ( $n = 3$  biological replicates). **e**, Representative H&E staining showing that  $35 \text{ mg kg}^{-1}$  cycloheximide treatment phenocopies rapamycin treatment 96 h after *Apc* deletion. Treatment began 24 h after induction ( $n = 3$  biological replicates). **f**, Representative IHC for BrdU showing a loss of proliferation in tumours after 72 h of  $35 \text{ mg kg}^{-1}$  cycloheximide treatment. ( $n = 3$  biological replicates). Arrow highlights normal proliferating crypts. Scale bar,  $100 \mu\text{m}$ .



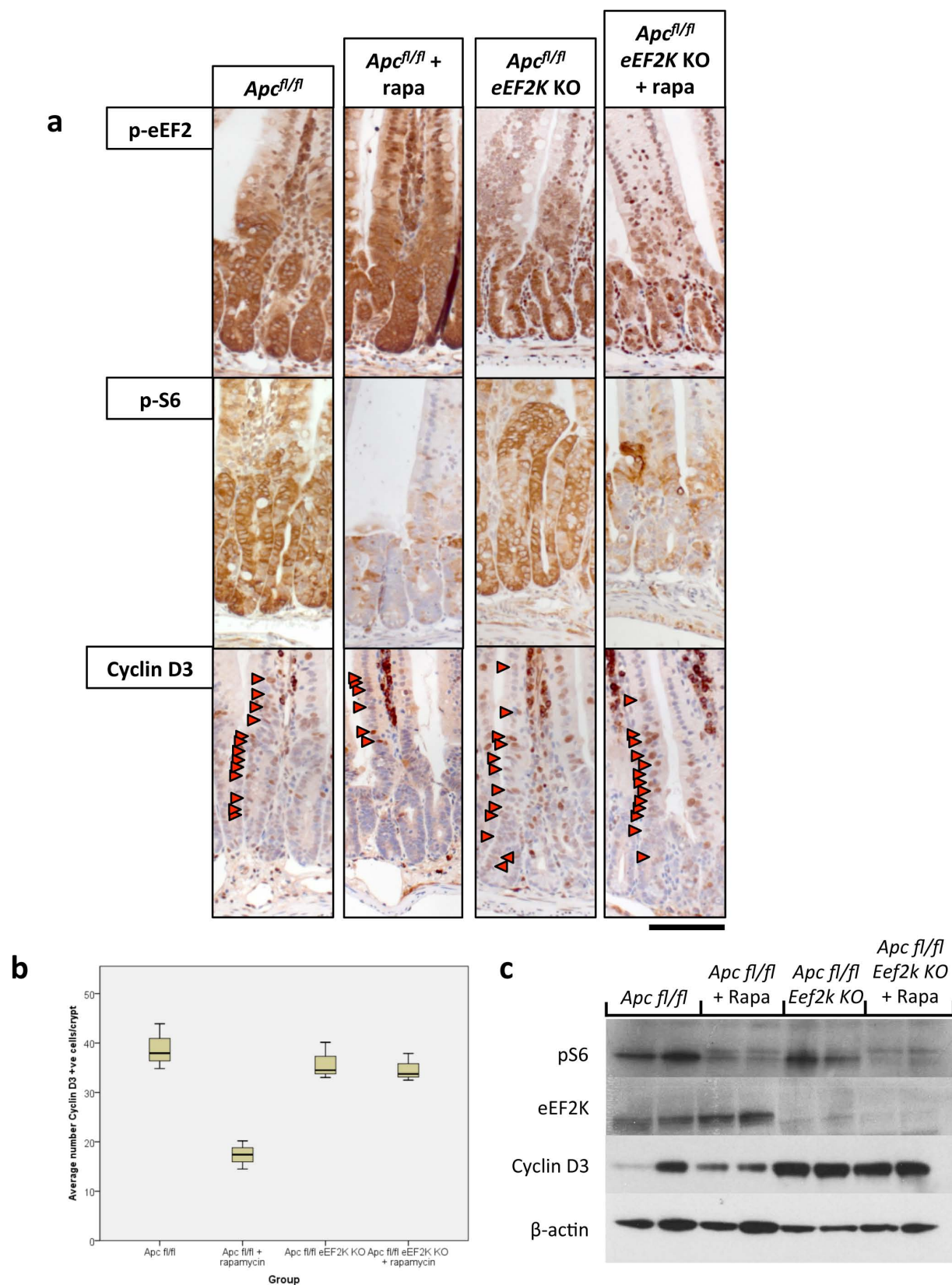
**Extended Data Figure 7 | *S6k* deletion decreases intestinal regeneration.** Graphical representation of findings, and boxplot showing that murine intestinal regeneration after irradiation is dependent on S6K. Animals were exposed to 14 Gy  $\gamma$ -irradiation, and intestinal regeneration was calculated 72 h after exposure by counting the number of viable crypts and multiplying that

by the average size of the regenerating crypts. Relative regeneration was calculated by comparing each group to wild-type regeneration. The rapamycin treatment arm is reproduced from Fig. 4 for visual clarity. Whiskers show maximum and minimum, black line shows median ( $n = 4$  per group). \* $P$  value = 0.034, Mann-Whitney  $U$  test.





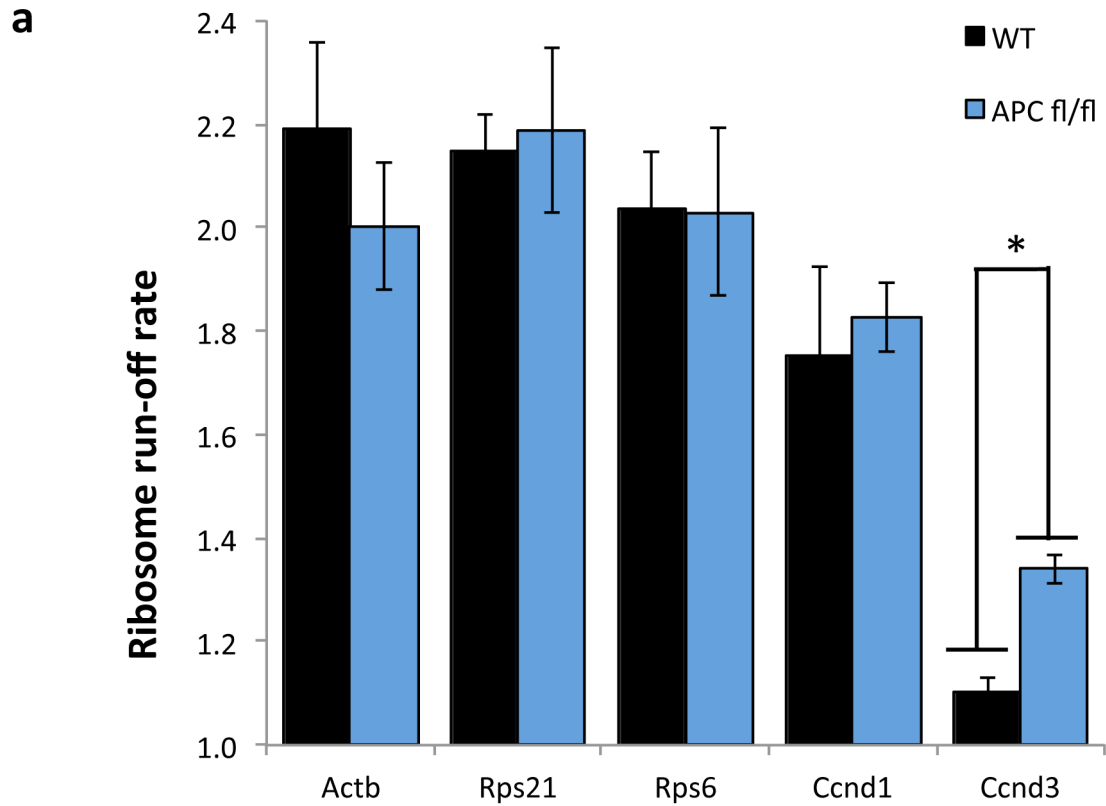
**Extended Data Figure 8 | *Eef2k* deletion drives resistance to rapamycin.**  
**a,** Representative IHC of phospho-eEF2 and phospho-RPS6 in wild-type (WT), *Apc*-deficient and *Apc*- and *Eef2k*-deficient mice (with or without 72 h 10 mg kg<sup>-1</sup> rapamycin (rapa) treatment) shows that rapamycin is unable to induce eEF2 phosphorylation in the absence of eEF2K (*n* = 6 biological replicates). KO, knockout. Scale bars, 100 μm.



### Extended Data Figure 9 | Cyclin D3 is regulated at the level of elongation.

**a**, Representative IHC of *Apc*-deleted intestines with or without *Eef2k* deletion. Antibodies to eEF2K, phospho-RPS6 and cyclin D3 are shown (representative of 3 biological replicates). After *Eef2k* knockout (KO), cyclin D3 levels are no longer decreased upon 10 mg kg<sup>-1</sup> rapamycin (rapa) treatment. **b**, Boxplot showing the number of cyclin-D3-positive cells per crypt, 96 h after *Apc* deletion, with and without 10 mg kg<sup>-1</sup> rapamycin treatment. Graph shows that

in *Eef2k* knockout animals, rapamycin no longer reduced cyclin D3 levels ( $n = 3$  biological replicates per group). \* $P$  value  $\leq 0.05$ , Mann-Whitney  $U$  test. **c**, Western blot analysis of intestinal epithelial cells from *Apc*-deleted and *Apc*-deleted *Eef2k* knockout, with and without 10 mg kg<sup>-1</sup> rapamycin. Antibodies to eEF2K, phospho-RPS6, cyclin D3 and  $\beta$ -actin are shown. Each well represents a different mouse from the relevant group. Cyclin D3 levels are no longer reduced after *Eef2k* deletion. Scale bar, 100  $\mu$ m.



**Extended Data Figure 10 | Ribosomes elongate faster on *Ccnd3* after *Apc* deletion.** The ribosome run-off rate of various messages was measured as in Fig. 3. Elongation of *Ccnd3* was significantly increased, while *Actb*, *Rps21*, *Rps6*

and *Ccnd1* remain unchanged. Data are average  $\pm$  s.e.m. ( $n = 3$  biological replicates per group). \* $P$  value  $\leq 0.05$ , Mann-Whitney  $U$  test.



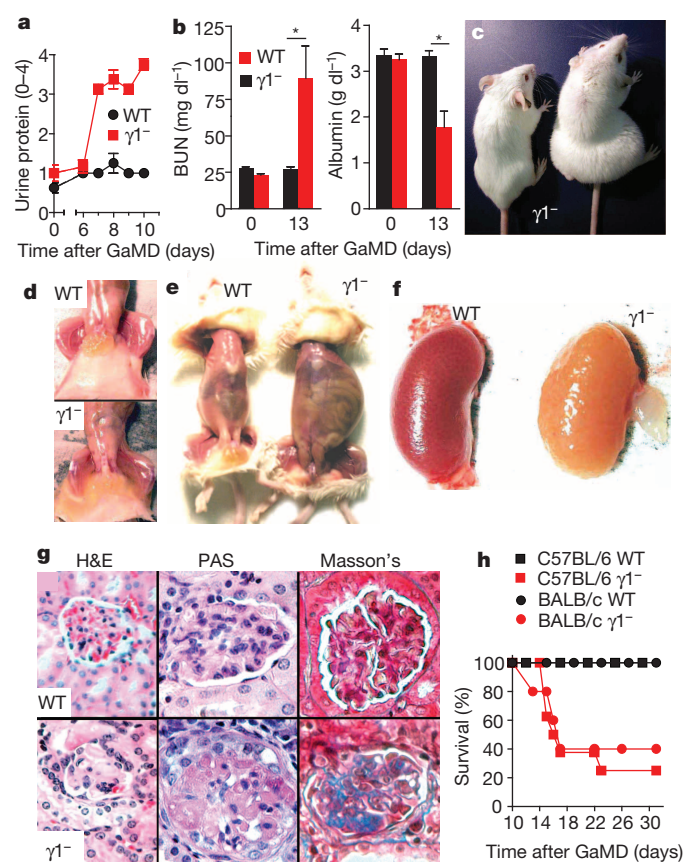
# IgG1 protects against renal disease in a mouse model of cryoglobulinaemia

Richard T. Strait<sup>1,2</sup>, Monica T. Posgai<sup>3</sup>, Ashley Mahler<sup>1</sup>, Nathaniel Barasa<sup>1</sup>, Chaim O. Jacob<sup>4</sup>, Jörg Köhl<sup>5,6</sup>, Marc Ehlers<sup>6</sup>, Keith Stringer<sup>7</sup>, Shiva Kumar Shanmukhappa<sup>7</sup>, David Witte<sup>7</sup>, Md Monir Hossain<sup>8</sup>, Marat Khodoun<sup>9</sup>, Andrew B. Herr<sup>3,5</sup> & Fred D. Finkelman<sup>5,9,10</sup>

Immunoglobulins protect against disease to a considerable extent by activating complement and stimulatory immunoglobulin crystallizable fragment receptors (Ig FcRs), and aggregating microbial pathogens<sup>1,2</sup>. Yet IgG1, the predominant murine serum Ig isotype, cannot activate complement by the classical pathway, binds more avidly to an inhibitory than to stimulatory FcRs, and has limited ability to aggregate pathogens<sup>1–3</sup>. In these regards, it resembles human IgG4 (ref. 4). We hypothesized that limited ability to activate effector mechanisms might protect against immune complex immunopathology. Here we show that IgG1-deficient ( $\gamma 1^{-/-}$ ) mice<sup>5</sup>, immunized with a potent antigen, develop lethal renal disease soon after they begin to produce antigen-specific antibody, whereas similarly immunized wild-type mice remain healthy. Surprisingly, renal disease in this model is complement and FcR independent and results from immune complex precipitation in glomerular capillaries, as in some cryoglobulinaemic humans<sup>6</sup>. IgG3, which self-associates to form large immune complexes<sup>7,8</sup>, accounts for more than 97% of the mouse Ig in this cryoglobulin; furthermore, glomerular disease develops when mice are injected with IgG3 anti-trinitrophenyl (TNP) monoclonal antibody followed by a TNP-labelled protein. Renal disease is prevented in both active and passive immunization models by antigen-specific IgG1; other isotypes are less potent at preventing disease. These observations demonstrate the adaptive significance of Ig isotypes that poorly activate effector mechanisms, reveal an immune-complex-dependent, complement- and FcR-independent nephrotoxic mechanism, and suggest that isotypes that poorly activate effector mechanisms may be useful for inhibiting immune complex immunopathology.

Immunization of wild-type BALB/c or C57BL/6 mice with a potent immunogen, goat anti-mouse IgD antiserum (GaMD), leads to a large, rapid, predominantly IgG1 antibody response to goat IgG (GIgG) and the generation of mouse IgG1–GIgG immune complexes<sup>9</sup>, but no noticeable disease. In contrast, GaMD-immunized  $\gamma 1^{-/-}$  BALB/c and C57BL/6 mice develop renal disease characterized by increased urinary protein, leukocyte esterase and erythrocytes (blood), starting 6–7 days after immunization, as well as increased blood concentration of urea (BUN) and decreased serum albumin, with anasarca (subcutaneous oedema) and peritoneal effusion (Fig. 1a–e and Extended Data Fig. 1a). Kidney colour in these mice changes from red/brown to yellow, reflecting dramatically decreased perfusion (Fig. 1f). Microscopically, glomerular capillaries contain IgG and complement deposits, but few inflammatory cells (Fig. 1g and Extended Data Fig. 1b, c). The microscopic damage is initially observed 6–7 days after GaMD immunization and is followed by disruption of glomerular structure and development of fibrosis (Fig. 1g and Extended Data Fig. 1c). Because no other organ damage was observed (data not shown), it is likely that renal insufficiency caused the death of 60–80% of  $\gamma 1^{-/-}$  mice by day 16–22 post-immunization (Fig. 1h).

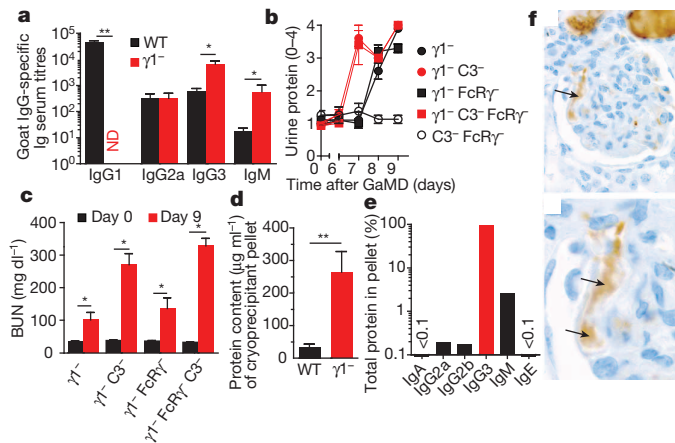
Lack of the normally dominant IgG1 response in GaMD-immunized  $\gamma 1^{-/-}$  mice was accompanied by increased production of IgG3, IgM and, in some experiments, IgG2a (Fig. 2a and Extended Data Fig. 2a). Because



**Figure 1 | GaMD-immunized  $\gamma 1^{-/-}$  mice develop lethal glomerulopathy.** a–h, Wild-type (WT) and  $\gamma 1^{-/-}$  mice (4 per group (a–g); or 8 or 10 per group (h)) were immunized with GaMD. **a**, Urine protein. **b**, Serum BUN and albumin. **c–g**, Representative photographs of mice 13 days after GaMD immunization demonstrate anasarca (**c**, **d**), ascites (**e**), kidney hypoperfusion (**f**) and glomerulopathy with PAS<sup>+</sup> deposits and fibrosis (blue colour on Masson's stain) (**g**) only in  $\gamma 1^{-/-}$  mice. H&E, haematoxylin and eosin; PAS, periodic acid–Schiff. Original magnification,  $\times 400$ . **h**, Survival curves. All figures show means  $\pm$  standard error of the mean (s.e.m.). Experiments depicted in all figures were repeated with similar results unless otherwise indicated. \* $P < 0.05$ , non-parametric Mann–Whitney two-tailed  $t$ -test.

<sup>1</sup>Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>2</sup>Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA. <sup>3</sup>Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA. <sup>4</sup>Department of Medicine, University of Southern California School of Medicine, Los Angeles, California 90033, USA. <sup>5</sup>Division of Immunobiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>6</sup>Institute for Systemic Inflammation Research, University of Lübeck, 23538 Lübeck, Germany. <sup>7</sup>Division of Pathology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>8</sup>Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>9</sup>Division of Immunology, Allergy and Rheumatology, Department of Medicine, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA. <sup>10</sup>Medical Service, Cincinnati Veterans Affairs Medical Center, Cincinnati, Ohio 45220, USA.





**Figure 2 | Glomerulopathy in GaMD-immunized  $\gamma 1^{-}$  mice is complement and FcR $\gamma$  independent and associated with IgG3 cryoglobulinaemia.**

**a**, Serum anti-goat IgG titres in wild-type (WT) and  $\gamma 1^{-}$  mice (4 per group) 8 days after GaMD immunization. **b**, **c**, Urine protein (**b**) and BUN (**c**) of GaMD-immunized  $\gamma 1^{-}$ ,  $\gamma 1^{-}/\text{FcR}\gamma^{-}$ ,  $\gamma 1^{-}/\text{C3}^{-}$ ,  $\text{C3}^{-}/\text{FcR}\gamma^{-}$  and  $\gamma 1^{-}/\text{C3}^{-}/\text{FcR}\gamma^{-}$  mice (5 per group). **d**, **e**, Serum cryoprecipitate protein and Ig isotype concentrations 6–7 days after GaMD immunization of wild-type and  $\gamma 1^{-}$  mice (7 or 8 per group). Only cryoprecipitates from  $\gamma 1^{-}$  mice contained detectable Ig. **f**, IgG3 (brown) in glomerular capillaries (arrows) of  $\gamma 1^{-}$  mice 8 days after GaMD (top, low magnification; bottom, high magnification). Original magnification,  $\times 400$ . \* $P < 0.05$ , \*\* $P < 0.005$ . ND, none detected. Error bars show s.e.m.

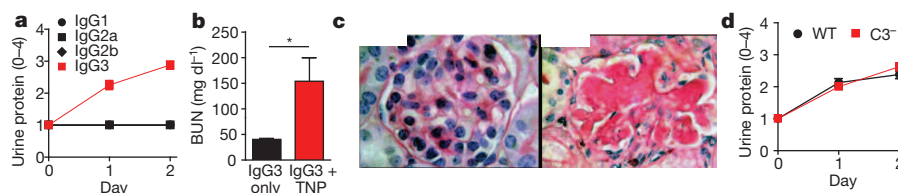
these isotypes, unlike IgG1, strongly activate complement, and IgG2a potently activates all stimulatory IgGFcRs<sup>1–3</sup>, we expected renal disease in  $\gamma 1^{-}$  mice to be complement and possibly FcR dependent. However, severe renal disease still developed in GaMD-immunized  $\gamma 1^{-}$  mice that lacked both C3—the complement component that is generally required for all complement activation pathways<sup>2</sup>—and FcR  $\gamma$ -chain (FcR $\gamma$ ), a required component of all stimulatory FcRs in mice<sup>10</sup> (Fig. 2b, c and Extended Data Fig. 2b). This was true even when these mice were also treated with C5a antagonists (Extended Data Fig. 3). Inhibition of IgG2a production with anti-interferon (IFN)- $\gamma$  monoclonal antibody<sup>11</sup> also failed to suppress kidney disease (Extended Data Fig. 2c, d). Additional studies eliminated the possibilities that renal disease in  $\gamma 1^{-}$  mice results from the persistence of circulating antigen or a decreased ratio of Ig to antigen that might form more inflammatory immune complexes (Extended Data Fig. 4).

These observations suggested that GaMD-induced kidney disease might be caused by a qualitative change in the immune complexes in immunized  $\gamma 1^{-}$  mice. Consistent with this, IgG3, the dominant isotype produced in these mice, generates large immune complexes by self-associating through Fc–Fc interactions<sup>7,8</sup>; these large immune complexes tend to reversibly precipitate at reduced temperature (that is, they are cryoglobulins) and at increased concentration (which occurs as plasma

undergoes ultrafiltration in glomeruli). Indeed, large cryoglobulin concentrations were found in plasma from GaMD-immunized  $\gamma 1^{-}$ , but not wild-type mice (Fig. 2d); cryoglobulin analysis demonstrated that IgG3 was the dominant mouse Ig constituent, although cryoglobulins also contained IgM (Fig. 2e). In keeping with this, deposits within glomerular capillaries were rich in IgG3 (Fig. 2f). A dominant role for IgM in this kidney disease model is unlikely because glomerular IgM, unlike glomerular IgG3, does not persist (Extended Data Fig. 5). Severe disease still develops in immunized mice deficient in both IgG1 and J chain (Extended Data Fig. 6), which produce little pentameric IgM<sup>12</sup>, and mice that lack activation-induced cytidine deaminase (AID) and consequently secrete only IgM do not develop kidney disease after GaMD immunization (data not shown).

A passive immunization model was used to further test the hypothesis that renal disease can be caused by IgG3–antigen immune complex precipitation in glomerular capillaries. Wild-type BALB/c mice were injected simultaneously with IgG3 anti-TNP monoclonal antibody intravenously and TNP-conjugated goat serum (TNP-GIgG) subcutaneously on days 0 and 1. These mice developed increased BUN, urine protein, leukocyte esterase and blood, and large deposits of amorphous material were observed in glomerular capillaries on day 2 (Fig. 3a–c and Extended Data Fig. 7a). Similar lesions developed in similarly treated C3-deficient mice (Fig. 3d and Extended Data Fig. 7b) and FcR $\gamma$ -deficient mice, as well as in C57BL/6 mice and in BALB/c wild-type mice when TNP-conjugated bovine serum albumin (TNP-BSA) was substituted for TNP-GIgG (data not shown). Wild-type mice injected with TNP-GIgG plus IgG1, IgG2a or IgG2b anti-TNP monoclonal antibody failed to develop renal disease (Fig. 3a and Extended Data Fig. 7a). None of the monoclonal antibodies induced disease when injected without TNP (Fig. 3b and data not shown).

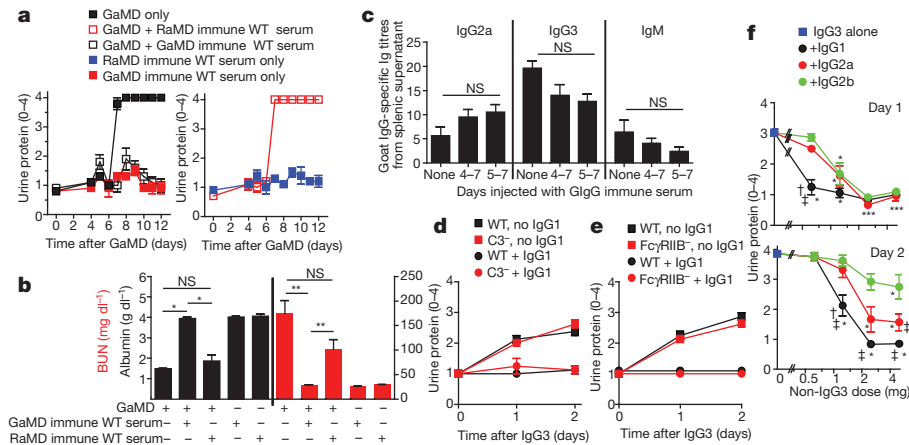
The unique pathogenicity of IgG3 raised the possibility that the other IgG isotypes might be able to inhibit IgG3-mediated disease. Consistent with this, GaMD induced only transient renal disease in  $\gamma 1^{+/-}$  mice, which produced  $\sim 50\%$  as much IgG1 as wild-type ( $\gamma 1^{+/+}$ ) mice, but similar amounts of IgG3 as  $\gamma 1^{-/-}$  mice (Extended Data Fig. 8). Similarly, development of proteinuria, hypoalbuminaemia and azotaemia in GaMD-immunized  $\gamma 1^{-/-}$  mice was suppressed by administration of the IgG1 anti-GIgG-rich serum from GaMD-immunized wild-type mice (GaMD immune WT serum). This suppression was antigen specific, because it was not observed with serum from rabbit anti-mouse IgD-immunized wild-type mice (RaMD immune WT serum) (Fig. 4a, b). Disease suppression by GaMD immune WT serum required initiation of treatment by day 5 after GaMD immunization (Extended Data Fig. 9a), when immunized mice first secrete IgG anti-GIgG. Importantly, injection of GaMD immune WT serum starting 4–5 days after GaMD immunization suppressed renal disease in  $\gamma 1^{-}$  mice without decreasing serum IgM, IgG2a or IgG3 levels and only modestly decreased production of any isotype by cultured spleen cells (Fig. 4c and Extended Data Fig. 9b). Thus, IgG1 primarily suppresses renal disease in our model by competing with IgG3 for antigen epitopes and/or changing the solubility of immune complexes rather than by decreasing IgG3 secretion; and the



**Figure 3 | Concurrent injection of wild-type mice with IgG3 anti-TNP monoclonal antibody and TNP-goat serum induces glomerulopathy.**

**a**, Wild-type (WT) mice (4 per group) were injected intravenously (i.v.) with 4 mg of mouse IgG1, IgG2a, IgG2b or IgG3 anti-TNP monoclonal antibody and subcutaneously (s.c.) with 100  $\mu\text{l}$  of TNP-goat serum on days 0 and 1. Urine protein was measured before injections and on days 1 and 2. **b**, **c**, Wild-type mice (4 per group) were injected with mouse IgG3 anti-TNP monoclonal

antibody plus or minus TNP-goat serum as in **a**. **b**, Day 2 mouse sera were analysed for BUN. **c**, Day 2 kidneys were stained with PAS (left, glomerulus from mouse that received only IgG3; right, glomerulus from mouse that received IgG3 plus TNP-goat serum). Representative of 3 mice per group. Original magnification,  $\times 400$ . **d**, Urine protein for BALB/c wild-type and C3<sup>−</sup> mice (4 per group) injected i.v. with 4 mg of IgG3 anti-TNP monoclonal antibody and s.c. with 400  $\mu\text{l}$  of TNP-goat serum on days 0 and 1. \* $P < 0.05$ .



**Figure 4 | Antigen-specific IgG1 prevents IgG3-mediated glomerulopathy.** a–f, BALB/c  $\gamma 1^{-}$  mice (5 per group) were injected with GaMD on day 0 and/or GaMD or rabbit anti-mouse IgD (RaMD) immune WT serum daily on days 4–7. a, b, Urine protein (a) and day 12 serum albumin and BUN levels (b). \* $P < 0.05$ , \*\* $P < 0.005$ . NS, not significant. c, BALB/c  $\gamma 1^{-}$  mice (5 per group) were injected with GaMD on day 0  $\pm$  GaMD immune WT serum daily on days 4–7 or 5–7. Spleens were harvested and cultured on day 8; 24 h culture supernatant IgG2a, IgG3 and IgM anti-goat IgG titres were determined. No anti-goat IgG antibody was detected in culture supernatants from non-immune spleen cells. d, Wild-type (WT) and C3-deficient (C3 $^{-}$ ) mice (4 per group) were injected s.c. with 100  $\mu$ l of TNP-goat serum and i.v. with 4 mg of mouse IgG3 anti-TNP plus or minus 5 mg of IgG1 anti-TNP monoclonal antibody on days 0 and 1. Urine was analysed on days 0, 1 and 2. e, Wild-type and Fc $\gamma$ RIIB-deficient (Fc $\gamma$ RIIB $^{-}$ ) mice (4 per group) were injected s.c. with 100  $\mu$ l of TNP-goat serum and i.v. with 4 mg of IgG3 anti-TNP plus or minus 5 mg of IgG1 anti-TNP on days 0 and 1. Urine was analysed on days 0, 1 and 2. f, BALB/c mice were injected i.v. with 4 mg IgG3 anti-TNP and s.c. with 1.4 mg of

TNP-BSA on days 0 and 1. Some mice were also injected with 0.625, 1.25, 2.5 or 5 mg of switch variant IgG1, IgG2a or IgG2b anti-TNP monoclonal antibodies on days 0 and 1. Urine protein was determined on day 0 (data not shown), day 1 (top) and day 2 (bottom). Results are pooled from a total of seven experiments. Group sizes were as follows. IgG3 alone: 19 mice; 0.625 mg of IgG1, IgG2a or IgG2b: 4 mice; 1.25 mg of IgG1, IgG2a or IgG2b: 8 mice; 2.5 mg of IgG1, IgG2a or IgG2b: 6 mice; 5 mg of IgG1, IgG2a or IgG2b: 8 or 9 mice. The significance of differences between treatment groups was determined with a normal regression model with the covariates: group; dose; and group–dose interaction. A *t*-test was used to evaluate the significance of the difference in least square means for each of these effects. This procedure was used to evaluate urine protein, measured at day 1 and day 2. A *P* value less than or equal to 0.05 after applying the Tukey adjustment for multiple comparisons was judged to be significant. \* $P < 0.05$  as compared to IgG3 alone; † $P < 0.05$  as compared to IgG2b plus IgG3; ‡ $P < 0.05$  as compared to IgG2a plus IgG3. Error bars show s.e.m.

increased IgG3 secretion by GaMD-immunized  $\gamma 1^{-}$  mice results from blocked isotype switching rather than from a lack of IgG1.

Consistent with our conclusion that IgG1 suppresses IgG3-induced renal disease by competing with IgG3 for antigen binding and/or changing immune complex solubility, IgG1 anti-TNP monoclonal antibody suppressed glomerular IgG3 deposition and disordered renal function when mice were injected with IgG3 anti-TNP monoclonal antibody plus TNP-BSA or TNP-goat serum (Fig. 4d, e and Extended Data Fig. 10). Very little IgG1 was found in the glomeruli of mice injected with both isotypes (Extended Data Fig. 10b), suggesting that the presence of IgG1 in an immune complex with IgG3 prevents glomerular immune complex deposition and/or increases its clearance. IgG1 suppression of IgG3-associated renal disease did not depend on C3 or Fc $\gamma$ RIIB (Fig. 4d, e and Extended Data Fig. 7b, c) and isotype control mouse IgG1 monoclonal antibody had no effect on IgG3-mediated disease (data not shown). IgG1 anti-TNP was more potent than IgG2a anti-TNP—and IgG2a anti-TNP was more potent than IgG2b anti-TNP—at preventing IgG3-mediated disease (Fig. 4f and Extended Data Fig. 7d), even though isotype switch variants of IgG1, IgG2a and IgG2b were used that had identical antigen-binding V regions<sup>13</sup>, avidity for TNP (Extended Data Fig. 7e) and non-specific binding to IgG3 (Extended Data Fig. 7f), and similar non-specific binding to themselves (Extended Data Fig. 7g). Preferential inhibition by IgG1 over IgG2a and IgG2b was also seen in studies with a second set of monoclonal antibodies that were not switch variants (data not shown).

The increased ability of IgG1 to inhibit IgG3-mediated renal disease may be a consequence of its short hinge region length and consequent low segmental flexibility. This may limit immune complex formation by decreasing the ability of IgG1 to bind bivalently to a ligand and by increasing the likelihood that it will sterically block binding of IgG3 (refs 14, 15), which could separate IgG3 molecules sufficiently to inhibit their self-association. Consistent with this possibility, IgG2a, which has a hinge region length and segmental flexibility that are intermediate between IgG1 and IgG2b<sup>14</sup>, had an intermediate ability to suppress

IgG3-mediated renal disease (Fig. 4f and Extended Data Fig. 7d). Thus, IgG1 may limit antibody-mediated disease in our model by suppressing the formation of immune complexes that become insoluble when they are concentrated by glomerular filtration. We cannot, however, eliminate the possibility that the addition of IgG1 to IgG3–antigen immune complexes facilitates their elimination by the reticulo-endothelial system, which could limit nephrotoxicity.

Our observations lead us to two important conclusions. First, we show that immune complexes can destroy kidney function by precipitating in glomerular capillaries, even in the absence of complement and FcR activation. The rapidity of capillary obstruction and the lack of an anaphylatoxin gradient in our model, as well as the ability of complement to increase immune complex solubility and elimination<sup>16</sup>, may explain the failure of complement to exacerbate disease despite its deposition in glomeruli. Second, we show that Ig isotypes that poorly activate effector mechanisms protect against disease caused by more proinflammatory isotypes. In this regard, isotypes such as mouse IgG1 and human IgG4 seem to act like partial agonists, which can cause immunopathology under some conditions, but prevent it by blocking the effects of other, more proinflammatory isotypes, in other circumstances. Indeed, functional similarities between mouse IgG1 and human IgG4 suggest that our observations in mice are applicable to humans. Mouse IgG1-mediated protection against immune complex deposition in our model is probably facilitated by its short hinge region, which limits antigen cross-linking by decreasing segmental flexibility<sup>14</sup>. Human IgG4 is likely to have an even greater ability to suppress immune complex development because, in addition to its short hinge region<sup>14</sup>, its labile inter-heavy-chain disulphide bond allows it to dissociate into univalent half molecules<sup>17</sup>.

Similarities between mouse IgG1 and human IgG4 in their ability to suppress disease caused by other isotypes may extend further. Although not shown in this paper, our preliminary observations, with additional collaborators, demonstrate that the absence of IgG1 promotes the development and severity of complement- and Fc $\gamma$ R-mediated diseases in

mice, including collagen-induced arthritis and experimental myasthenia gravis. Thus, mouse IgG1 probably suppresses disease mediated by complement and FcγRs, as well as disease mediated by excessive intravascular formation of insoluble immune complexes. The inability of human IgG4 to activate complement<sup>2,3</sup> and its poor binding to FcγRs<sup>4</sup> suggest that it can similarly limit organ damage in complement- and FcγR-mediated diseases. These considerations raise the possibility of using human IgG4 antibodies to suppress autoimmune and immune complex disorders that are mediated by other isotypes, an approach that might be amplified by making IgG4 antibodies even more immunosuppressive by increasing their sialylation<sup>18</sup>, galactosylation<sup>19</sup> and/or affinity for FcγRIIB<sup>20</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 28 April 2011; accepted 16 September 2014.**

**Published online 2 November 2014; corrected online 21 January 2015 (see full-text HTML version for details).**

1. Nimmerjahn, F. & Ravetch, J. V. Fcγ receptors as regulators of immune responses. *Nature Rev. Immunol.* **8**, 34–47 (2008).
2. Barrington, R., Zhang, M., Fischer, M. & Carroll, M. C. The role of complement in inflammation and adaptive immunity. *Immunol. Rev.* **180**, 5–15 (2001).
3. Snapper, C. & Finkelman, F. In *Fundamental Immunology* (ed. Paul, W. E.) 837–863 (Lippincott-Raven, 1999).
4. Bruhns, P. *et al.* Specificity and affinity of human Fcγ receptors and their polymorphic variants for human IgG subclasses. *Blood* **113**, 3716–3725 (2009).
5. Jung, S., Rajewsky, K. & Radbruch, A. Shutdown of class switch recombination by deletion of a switch region control element. *Science* **259**, 984–987 (1993).
6. Roccatello, D. *et al.* Multicenter study on hepatitis C virus-related cryoglobulinemic glomerulonephritis. *Am. J. Kidney Dis.* **49**, 69–82 (2007).
7. Nishimura, Y. & Nakamura, H. Human monoclonal cryoimmunoglobulins. I. Molecular properties of IgG3κ (Jir protein) and the cryo-coprecipitability of its molecular fragments by papain. *J. Biochem.* **95**, 255–265 (1984).
8. Abdelmoula, M. *et al.* IgG3 is the major source of cryoglobulins in mice. *J. Immunol.* **143**, 526–532 (1989).
9. Finkelman, F. D. *et al.* Polyclonal activation of the murine immune system by an antibody to IgD. II. Generation of polyclonal antibody production and cells with surface IgG. *J. Immunol.* **129**, 638–646 (1982).
10. Ravetch, J. V. & Bolland, S. IgG Fc receptors. *Annu. Rev. Immunol.* **19**, 275–290 (2001).
11. Finkelman, F. D., Katona, I. M., Mosmann, T. R. & Coffman, R. L. IFN-γ regulates the isotypes of Ig secreted during *in vivo* humoral immune responses. *J. Immunol.* **140**, 1022–1027 (1988).
12. Erlandsson, L., Andersson, K., Sigvardsson, M., Lycke, N. & Leanderson, T. Mice with an inactivated joining chain locus have perturbed IgM secretion. *Eur. J. Immunol.* **28**, 2355–2365 (1998).
13. Boot, J. H., Geerts, M. E., De Groot, E. R. & Aarden, L. A. Murine monoclonal isotype switch variants. Detection with rat monoclonal antibodies in ELISA and isolation by sequential sublining. *J. Immunol. Methods* **106**, 195–202 (1988).
14. Dangi, J. L. *et al.* Segmental flexibility and complement fixation of genetically engineered chimeric human, rabbit and mouse antibodies. *EMBO J.* **7**, 1989–1994 (1988).
15. Schneider, W. P., Wensel, T. G., Stryer, L. & Oi, V. T. Genetically engineered immunoglobulins reveal structural features controlling segmental flexibility. *Proc. Natl Acad. Sci. USA* **85**, 2509–2513 (1988).
16. Sheerin, N. S., Springall, T., Abe, K. & Sacks, S. H. Protection and injury: the differing roles of complement in the development of glomerular injury. *Eur. J. Immunol.* **31**, 1255–1260 (2001).
17. van der Neut Kolfschoten, M. *et al.* Anti-inflammatory activity of human IgG4 antibodies by dynamic Fab arm exchange. *Science* **317**, 1554–1557 (2007).
18. Anthony, R. M., Kobayashi, T., Wermeling, F. & Ravetch, J. V. Intravenous gammaglobulin suppresses inflammation through a novel T<sub>H</sub>2 pathway. *Nature* **475**, 110–113 (2011).
19. Karsten, C. M. *et al.* Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. *Nature Med.* **18**, 1401–1406 (2012).
20. Chu, S. Y. *et al.* Reduction of total IgE by targeted coengagement of IgE B-cell receptor and FcγRIIB with Fc-engineered antibody. *J. Allergy Clin. Immunol.* **129**, 1102–1115 (2012).

**Acknowledgements** We thank M. Wills-Karp for C3-deficient mice, D. Metzger for J-chain-deficient mice, M. Muramatsu for AID-deficient mice, S. Izui for a hybridoma that secretes mouse IgG1 anti-TNP monoclonal antibody, M. Robson and L. Aarden for switch variant hybridomas that secrete mouse IgG1, IgG2a and IgG2b anti-TNP monoclonal antibodies, B. DiPasquale for histological staining, and J. Lambris for C5aR antagonist. Research was supported by a US Department of Veterans Affairs Merit Award, National Institutes of Health R01 AI072040, and the University of Cincinnati and Cincinnati Children's Hospital, all to F.D.F.

**Author Contributions** R.T.S. planned, performed, interpreted and directed all experiments and participated in all drafts of the manuscript; M.K., A.M. and N.B. performed experiments; C.O.J. provided mice; J.K. provided mice and reagents and participated in planning of all complement-related studies; M.E. participated in the planning and interpretation of isotype switch variant studies; S.K.S., K.S. and D.W. performed and interpreted histological studies; M.T.P. and A.B.H. contributed to investigation of the importance of antibody hinge region length; M.M.H. performed statistical analyses; and F.D.F. designed the study, analysed data and wrote the initial draft of the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.D.F. (finkelfd@UCMAIL.UC.EDU).



## METHODS

**Mice.** All mice were bred and maintained in the specific pathogen-free (SPF) facility at the Cincinnati Children's Research Foundation and all experiments were done with the approval of and in accordance with regulatory guidelines and standards set by the Institutional Animal Care and Use Committee of Cincinnati Children's Hospital Medical Center. Male and female mice were used between the ages of 8 and 20 weeks. BALB/c and/or C57BL/6 background  $\gamma 1$ -deficient mice<sup>2</sup>, FcR $\gamma$ -deficient (Taconic), C3-deficient mice (a gift from M. Wills-Karp), J-chain deficient mice (a gift from D. Metzger), Fc $\gamma$ RIIB-deficient, AID-deficient (a gift from M. Muramatsu)<sup>21</sup> and wild-type control mice were bred in our colony. Double and triple gene-deficient mice, made by crossbreeding the single gene-deficient mice, were typed by PCR. Typing for  $\gamma 1$  and C3 deficiency was confirmed by gel double diffusion assay of serum. Wild-type littermates of the double and triple gene-deficient mice were used as controls. Mice of the appropriate genotype were randomly assigned to groups, but a specific randomization program was not used. Studies were not blinded.

**Active model for induction of immune complex renal disease.** Mice were injected s.c. with 0.2 ml (BALB/c) or 0.4 ml (C57BL/6) of GaMD on day 0. In some experiments mice were also injected intraperitoneally (i.p.) daily with pooled day 12 serum from GaMD or RaMD immunized WT mice (GaMD immune WT serum and RaMD immune WT serum, respectively). Spontaneously micturated urine was collected from mice on specific days and assayed for protein, leukocyte esterase and blood content by urine dipstick. Serum was collected by tail vein puncture and kidneys were preserved in formalin or glutaraldehyde or frozen in OCT for histological evaluation.

**Passive model for immune complex induction of renal disease.** Mice were simultaneously injected with mouse IgG3 anti-TNP monoclonal antibody i.v. and TNP-goat serum or TNP-BSA s.c. on days 0 and 1. Some mice also received mouse IgG1, IgG2a or IgG2b anti-TNP monoclonal antibody isotype switch variants or in some cases non-switch variant monoclonal antibodies i.v. on days 0 and 1.

**Reagents.** Hybridomas were obtained from the following sources: 9A6 (mouse IgG3 anti-TNP monoclonal antibody), a gift from S. Izui<sup>22</sup>; 1B7.11 (mouse IgG1 anti-TNP monoclonal antibody), from the American Type Culture Collection; HY1.2 and C1040 (mouse IgG2a anti-TNP monoclonal antibodies) and GORK (mouse IgG2b anti-TNP monoclonal antibody), a gift from B. Heyman; switch variant mouse IgG1, IgG2a and IgG2b anti-TNP monoclonal antibodies, a gift from M. Robson<sup>23</sup>; and XM6-6 (rat IgG1 anti-mouse IFN- $\gamma$ )<sup>24</sup>, from DNAX. A hybridoma that secretes mouse IgG1 anti-FITC monoclonal antibody was produced in house. Hybridomas were grown as ascites in pristane-primed athymic nude mice and monoclonal antibodies were purified from ascites by ammonium sulphate precipitation (25–50% for all IgGs except 20–30% for IgG3), followed by DE-52 cation exchange chromatography for the IgG isotypes. IgG2a anti-TNP monoclonal antibody was also purchased from Bio X Cell. GaMD and RaMD antisera were made as described<sup>25</sup>. Mouse hyperimmune antisera to goat (GaMD immune WT serum) and rabbit (RaMD immune WT serum) were made by injecting wild-type mice s.c. every 14 days for several injections with GaMD or RaMD, respectively, and pooling serum collected 10–12 days after each immunization. Non-immune mouse serum was collected by tail vein bleeding from untreated BALB/c or C57BL/6 wild-type mice and pooled. The C5aR antagonists JPE 1275 (ref. 26) (a gift from J. Lambris) and A8871-73 (ref. 27) were injected at 20  $\mu$ g or 10  $\mu$ M doses, respectively, i.p. every 12 h starting on day 4 and ending on day 8 after GaMD immunization.

**Urine and serum measurements.** Urine protein, leukocyte esterase and erythrocyte levels were measured on fresh, freely excreted urine using Multistix 10 (Becton-Dickson). Measurements were on a colorimetric scale, ranging from 0 to 4 for protein and 0 to 3 for leukocyte esterase and erythrocytes, as per the manufacturer's instructions. Serum albumin and BUN were measured using a Beckman DXC courtesy of R. Louderbeck, Veterans Administration Medical Center, Cincinnati, Ohio. Serum, splenic supernatants and re-suspended cryoglobulin pellets were analysed for total and GIG-specific IgG1, IgG2a/c, IgG2b, IgG3, IgE, IgA and IgM content, using standard sandwich ELISA with paired anti-Ig isotype monoclonal antibodies for each Ig isotype (BD-Pharmingen and eBioscience). Sera, splenic supernatants and re-suspended cryoglobulin solutions were titred for GIG-specific antibody levels by ELISA, as previously described<sup>25</sup>, using wells coated with 5  $\mu$ g ml<sup>-1</sup> of goat IgG. Gel double diffusion was used to identify mouse IgG1, mouse C3 and GIG in mouse serum with antibodies purchased from Bethyl.

**TNP-goat serum and TNP-BSA.** Goat serum or BSA were conjugated to TNP as previously described<sup>28</sup>.

**Immunofluorescence microscopy.** Kidneys were harvested from mice and immediately placed in OCT and snap frozen in liquid nitrogen. OCT-embedded kidneys were stored at  $-80^{\circ}\text{C}$ . Frozen tissue sections were cut, mounted on glass slides, fixed in acetone and air-dried. After rehydration and blocking, immunofluorescence microscopy was performed with FITC-labelled anti-C3 and anti-mouse IgG antibodies (ICN Biomedicals/Cappel). After washing, coverslips were applied to slides after addition of anti-fade medium that contained 4',6-diamidino-2-phenylindole (DAPI)

(Prolong Gold; Invitrogen). Slides were assessed microscopically and photographed at an original magnification of  $\times 400$  using an RT Slider digital camera (Diagnostic Instruments) mounted on an E600 fluorescent microscope (Nikon Instruments).

**Immunostaining microscopy.** Kidneys were harvested from mice and immediately placed in formalin for a minimum of 3 days before embedding in paraffin. For IgG3 staining, de-paraffinized sections were incubated with goat anti-mouse IgG3 antibody (Jackson ImmunoResearch Laboratories) for 1 h, then incubated for 12 min with biotin-donkey anti-goat IgG antibody (Jackson ImmunoResearch Laboratories) and visualized with an iVIEW Plus Detection Kit (Ventana Medical Systems). Staining for IgG1 was performed with a rabbit anti-mouse IgG1 antibody (Novus Biologicals), followed by biotin-donkey anti-rabbit antibody (Jackson ImmunoResearch Laboratories). All antibodies were used at 1:100 dilutions and all staining was performed with a Discovery XT (Ventana Medical Systems).

**Quantitation of splenic Ig production.** Spleens were diced into 1–2 mm cubes, washed in ice-cold PBS, then transferred to a Petri dish with 5 ml of RPMI medium 1640 supplemented with fetal bovine serum, penicillin, gentamicin, HEPES, sodium pyruvate, essential amino acids and 2-mercaptoethanol and cultured at  $37^{\circ}\text{C}$ , 5% CO<sub>2</sub> for 24 h. Supernatants were collected, separated into aliquots, frozen and stored at  $-80^{\circ}\text{C}$  until analysed.

**Cryoprecipitant collection.** Blood was collected in a polystyrene centrifuge tube and immediately incubated for 4 h at  $37^{\circ}\text{C}$ . The tube was centrifuged and solid material was removed. Sera were incubated at  $4^{\circ}\text{C}$  for 7 days. Precipitates were obtained by centrifugation, washed 3 times in ice-cold saline, then re-suspended in a volume of saline equal to that of the initial serum sample and heated to  $37^{\circ}\text{C}$  for 2 h to dissolve cryoprecipitates. Samples were then re-centrifuged at room temperature and supernatants were collected.

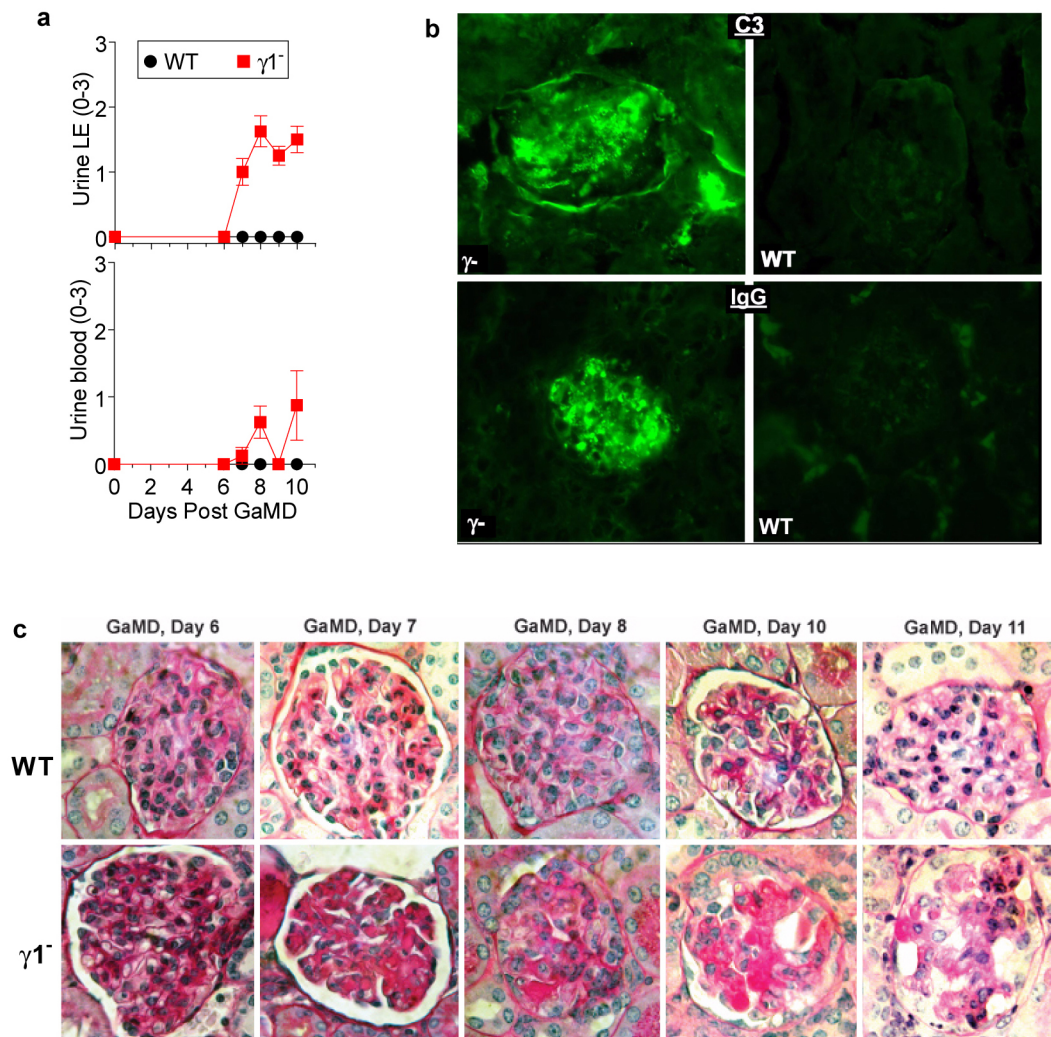
**Anti-TNP monoclonal antibody ELISA.** For avidity measurement ELISA plates were coated with 10  $\mu$ g TNP-OVA per ml overnight followed by varying concentrations of the monoclonal antibodies. This was followed by horseradish peroxidase (HRP)- or biotin-labelled anti-mouse Ig isotype-specific monoclonal antibody purchased from BD bioscience, streptavidin-HRP if needed and SuperSignal ELISA substrate from Pierce. For IgG3 binding or self-association measurement, ELISA plates were coated with 10  $\mu$ g mouse IgG1, IgG2a, IgG2b or IgG3 anti-TNP monoclonal antibody per ml overnight followed by varying concentrations of biotin-labelled mouse IgG1, IgG2a, IgG2b or IgG3 anti-TNP monoclonal antibodies. This was followed by streptavidin-HRP and SuperSignal substrate from Pierce.

**Statistics.** The non-parametric Mann-Whitney two-tailed *t*-test (GraphPad Prism 5.0; GraphPad Software) was used to compare Ig levels, BUN and albumin concentrations between different groups of mice. A *P* value  $< 0.05$  was considered significant. A more complex test was used to compare the multiple samples in Fig. 4f and Extended Data Fig. 7d (see figure legends for details).

Sample size was calculated with a tool for comparing two independent samples on the website <http://www.stat.ubc.ca/~rollin/stats/ssize/n2.html>. Sample size calculations for initial studies were based on the assumptions that a one-sided test would be used to test the hypothesis that the mean for the normal (usually wild-type) group would be three times as large (or one-third as large) as the mean for the abnormal (usually  $\gamma 1$ -deficient) group, that the common standard deviation would be the same size as the smaller mean, that the type I error rate would be 0.05 and that the desired power would be 0.80. These assumptions suggested a sample size of 4 for each sample. In practice, we often used the results of our initial studies to determine whether this sample size was sufficient to yield the desired power or indicate the larger sample size that would yield a significant result if the trend observed in the initial study predicted the results obtained with the larger sample size.

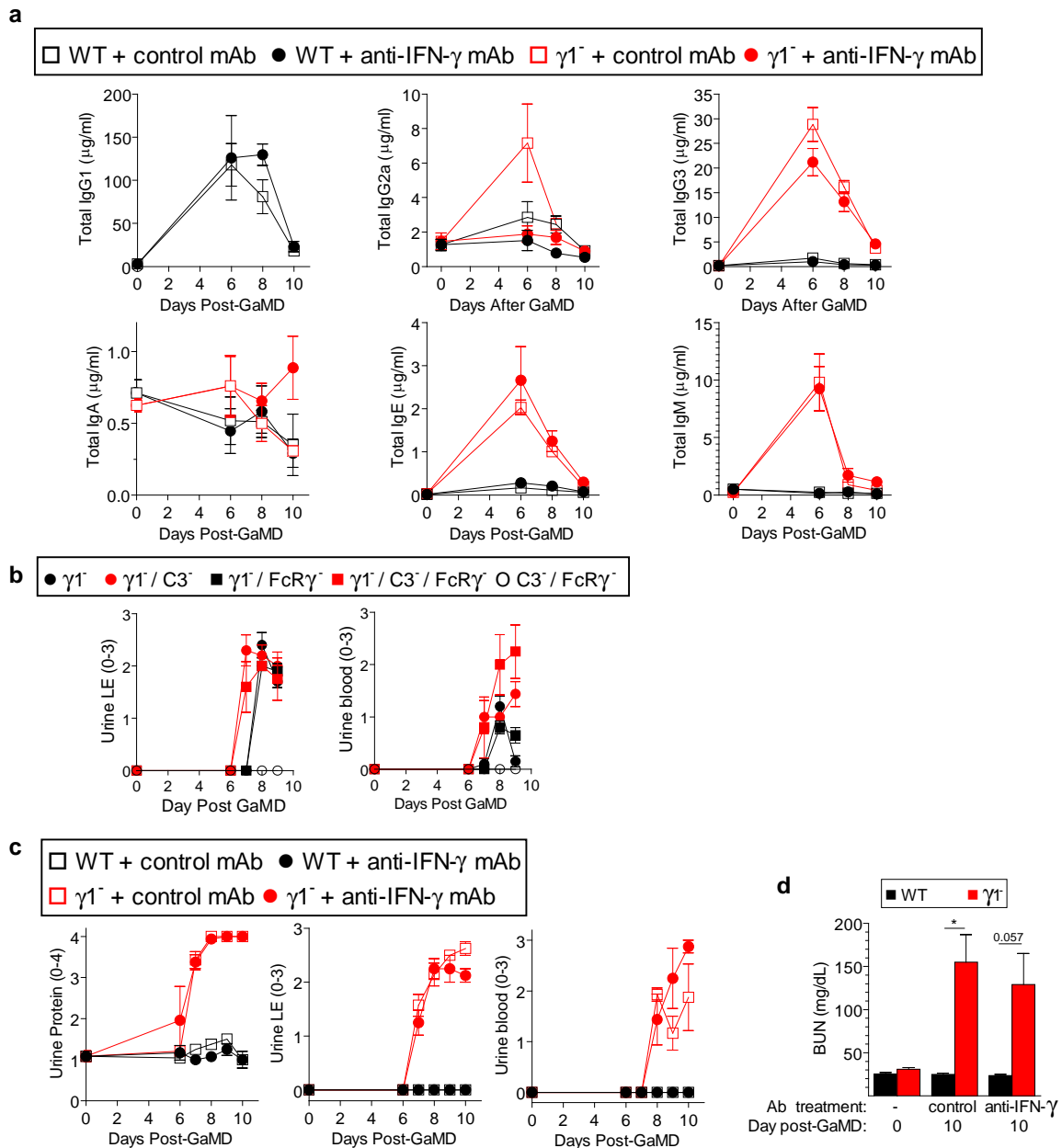
- Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
- Fulpius, T. *et al.* Glomerulopathy induced by IgG3 anti-trinitrophenyl monoclonal cryoglobulins derived from non-autoimmune mice. *Kidney Int.* **45**, 962–971 (1994).
- Giorgini, A. *et al.* Fc $\gamma$ RIII and Fc $\gamma$ RIV are indispensable for acute glomerular inflammation induced by switch variant monoclonal antibodies. *J. Immunol.* **181**, 8745–8752 (2008).
- Finkelman, F. D., Katona, I. M., Mosmann, T. R. & Coffman, R. L. IFN- $\gamma$  regulates the isotypes of Ig secreted during *in vivo* humoral immune responses. *J. Immunol.* **140**, 1022–1027 (1988).
- Finkelman, F. D. *et al.* Polyclonal activation of the murine immune system by an antibody to IgD. II. Generation of polyclonal antibody production and cells with surface IgG. *J. Immunol.* **129**, 638–646 (1982).
- Boor, P. *et al.* Complement C5 mediates experimental tubulointerstitial fibrosis. *J. Am. Soc. Nephrol.* **18**, 1508–1515 (2007).
- Köhl, J. *et al.* A regulatory role for the C5a anaphylatoxin in type 2 immunity in asthma. *J. Clin. Invest.* **116**, 783–796 (2006).
- Strait, R. T., Morris, S. C. & Finkelman, F. D. IgG-blocking antibodies inhibit IgE-mediated anaphylaxis *in vivo* through both antigen interception and Fc $\gamma$ RIIb cross-linking. *J. Clin. Invest.* **116**, 833–841 (2006).





**Extended Data Figure 1 | GaMD immunization of  $\gamma 1^{-}$  mice induces renal dysfunction and glomerular deposition of PAS<sup>+</sup> material that includes IgG and complement.** **a**, Wild-type (WT) and  $\gamma 1^{-}$  mice (4 per group) were immunized with GaMD. Urine leukocyte esterase and blood were obtained. **b**, Representative photomicrographs of glomeruli stained for C3 (top) or total mouse IgG (bottom) from wild-type (right) and  $\gamma 1^{-/-}$  mice (left), 12 days

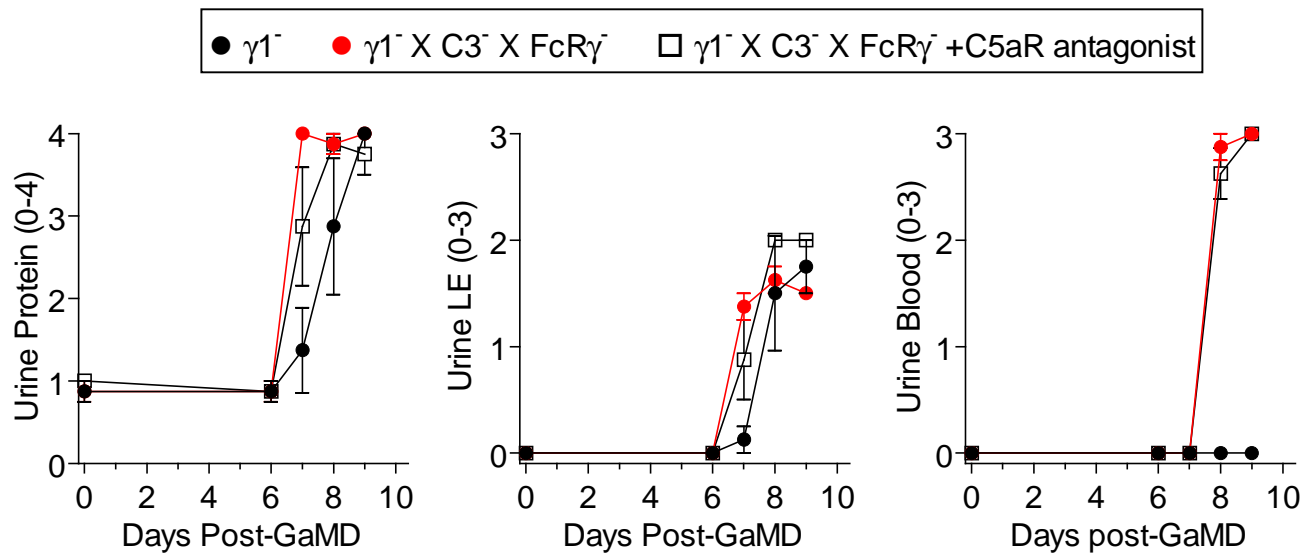
after GaMD immunization (3 mice per group). **c**, Deposition of amorphous PAS<sup>+</sup> material in glomeruli of  $\gamma 1^{-}$ , but not wild type begins ~7 days after GaMD immunization and leads to glomerular destruction by day 9. Note the scarcity of inflammatory cells in glomeruli. Representative data from 6 mice per group. Original magnification,  $\times 400$  for **b**,  $\times 200$  for **c**.



**Extended Data Figure 2 | The development of kidney disease in GaMD-immunized  $\gamma 1^{-}$  mice is independent of IFN- $\gamma$ , IgG2a, C3 and FcR $\gamma$ .**

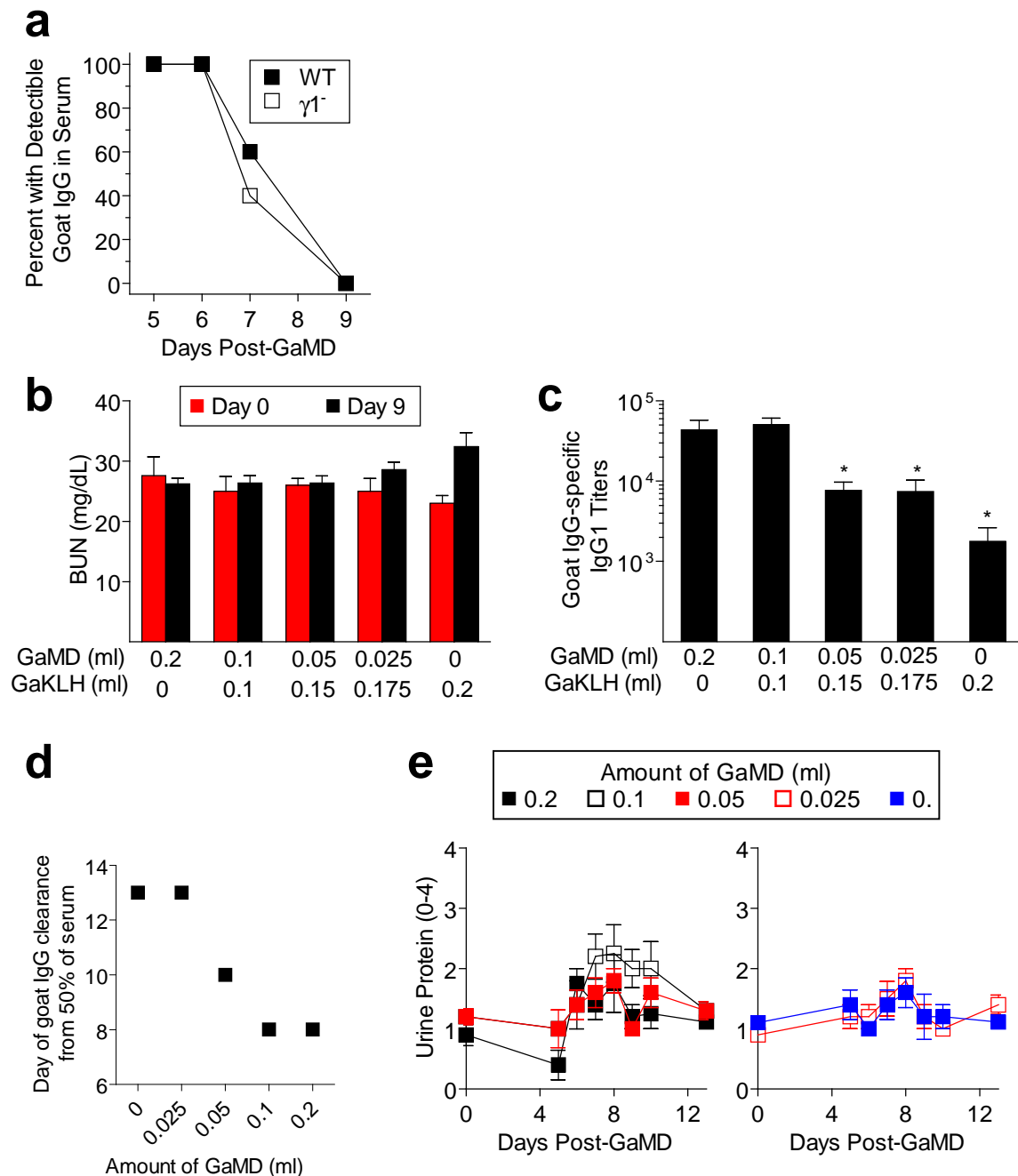
**a–d**, BALB/c wild-type (WT) and  $\gamma 1^{-}$  mice (5 per group) were immunized with GaMD on day 0 and injected with 1 mg of either anti-IFN- $\gamma$  or control monoclonal antibody (mAb) on days 0 and 5. **a**, Total levels of all Ig isotypes were determined in 24 h culture supernatants of spleen cells harvested on the days shown. **b**, GaMD-immunized  $\gamma 1^{-}$ ,  $\gamma 1^{-}/FcR\gamma^{-}$ ,  $\gamma 1^{-}/C3^{-}$ ,  $C3^{-}/FcR\gamma^{-}$

and  $\gamma 1^{-}/C3^{-}/FcR\gamma^{-}$  mice (5 per group) had their urine tested for leukocyte esterase and blood on the days shown. **c**, **d**, BALB/c wild-type and  $\gamma 1^{-}$  mice (5 per group) were immunized with GaMD on day 0 and injected with 1 mg of either anti-IFN- $\gamma$  or control monoclonal antibody on days 0 and 5. **c**, Urine obtained on days indicated was assayed for protein, leukocyte esterase and blood. **d**, BUN levels were determined before and 10 days after GaMD immunization. \* $P < 0.05$ .



**Extended Data Figure 3 | Neither complement nor stimulatory FcRs are required for renal disease development in GaMD-immunized  $\gamma 1^{-/-}$  mice.** Mixed background  $\gamma 1^{-/-}$  and  $\gamma 1^{-/-} \times C3^{-/-} \times FcR\gamma^{-/-}$  mice (4 per group) were

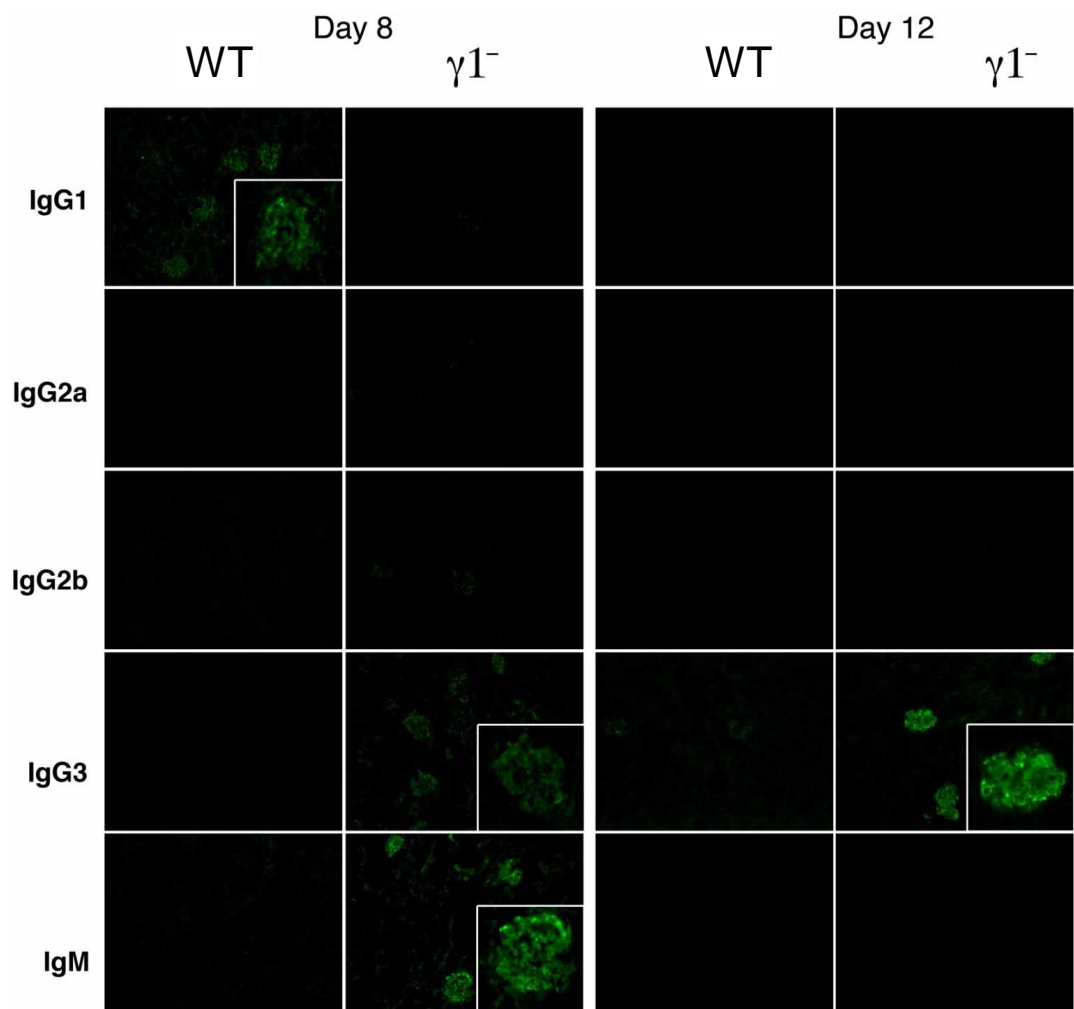
immunized with GaMD plus or minus C5aR antagonist. Urinalyses were obtained at baseline and daily starting on day 6. LE, leukocyte esterase.



**Extended Data Figure 4 | Delayed antigen elimination does not account for renal disease in GaMD-immunized  $\gamma 1^{-}$  mice.** **a**, BALB/c wild-type (WT) and  $\gamma 1^{-}$  mice (10 per group) were immunized s.c. with GaMD. Sera obtained 5, 6, 7 and 9 days later were evaluated by gel double diffusion for the presence of goat IgG. **b–e**, BALB/c wild-type mice (4 or 5 per group) were injected s.c. with a total of 0.2 ml of different mixtures of GaMD and goat anti-KLH

antisera. **b**, **c**, Mouse sera collected 9 days later were assayed for BUN (**b**) and IgG1 anti-goat IgG antibody (**c**). **d**, Sera obtained 6–13 days post-immunization were evaluated by gel double diffusion for the presence of goat IgG. **e**, Urine samples collected 4–12 days post-immunization were analysed for protein. \* $P < 0.05$ , \*\* $P < 0.005$ .

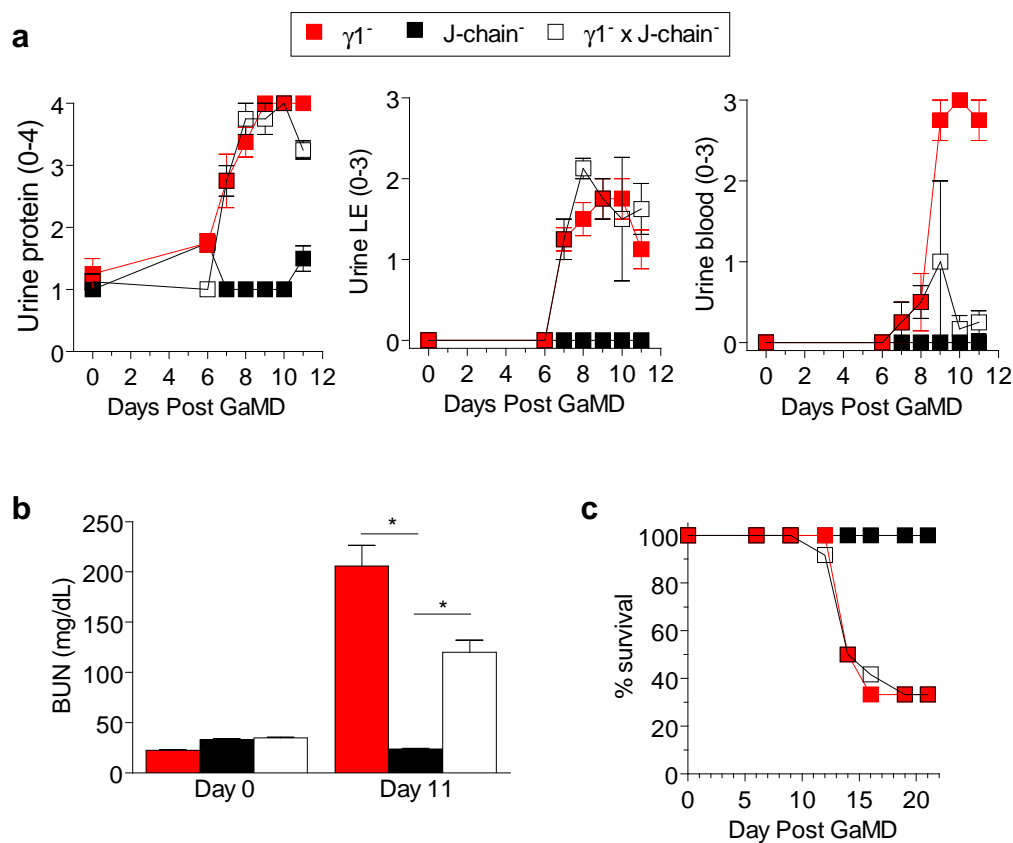




\*\*No staining for any immunoglobulin was seen in the kidney for unimmunized mice

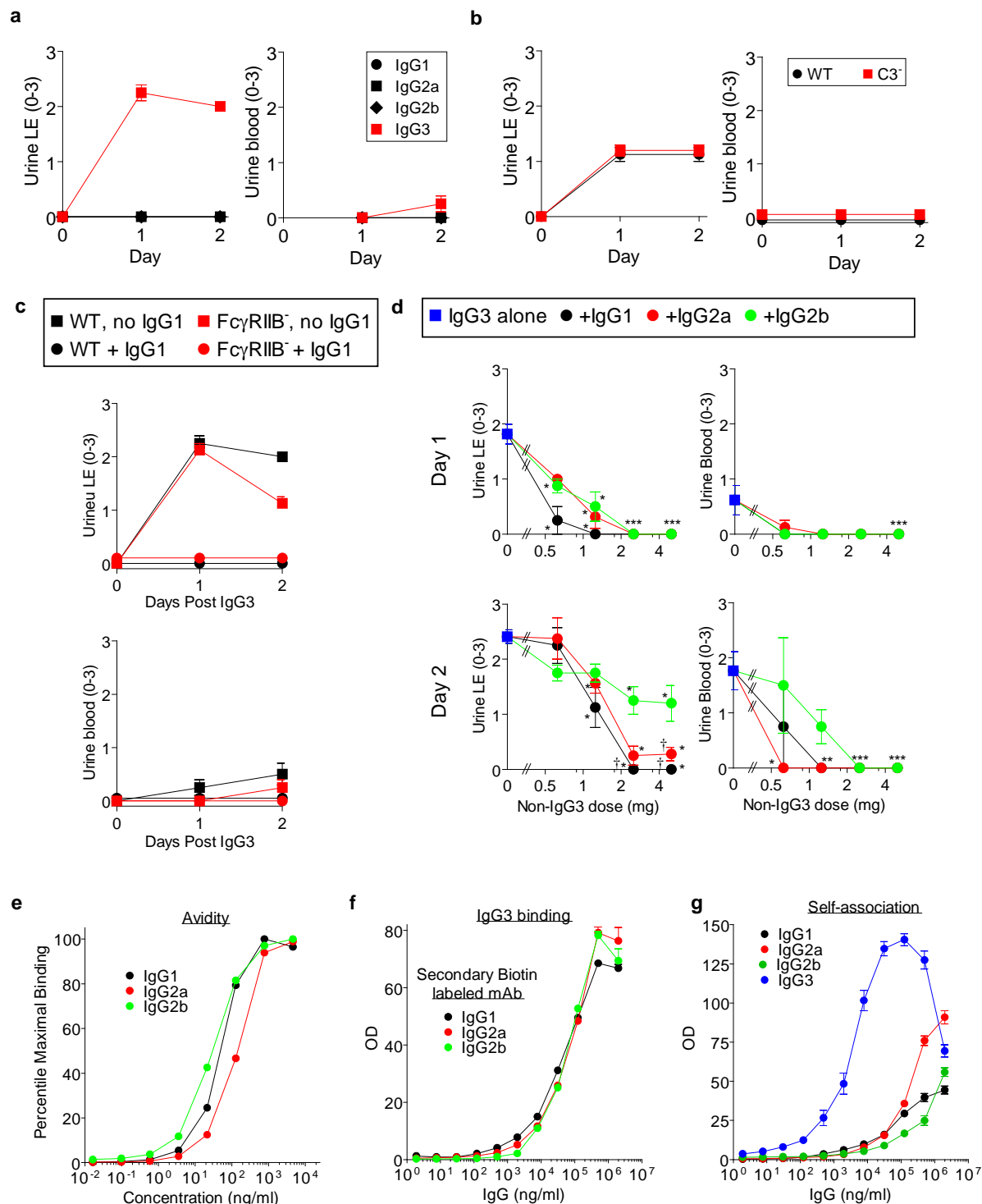
**Extended Data Figure 5 | IgG3 immune complexes persist and accumulate in the glomeruli of GaMD-immunized  $\gamma 1^{-}$  mice.** BALB/c wild-type (WT) and  $\gamma 1^{-}$  mice were left untreated or were immunized with GaMD. Kidney sections were stained for mouse IgG1, IgG2a, IgG2b, IgG3 and IgM and

12 days after immunization. Representative photomicrographs from three GaMD-immunized mice are shown. Insets show magnified views. No staining was observed with sections from unimmunized mice (data not shown). Original magnification,  $\times 400$ .



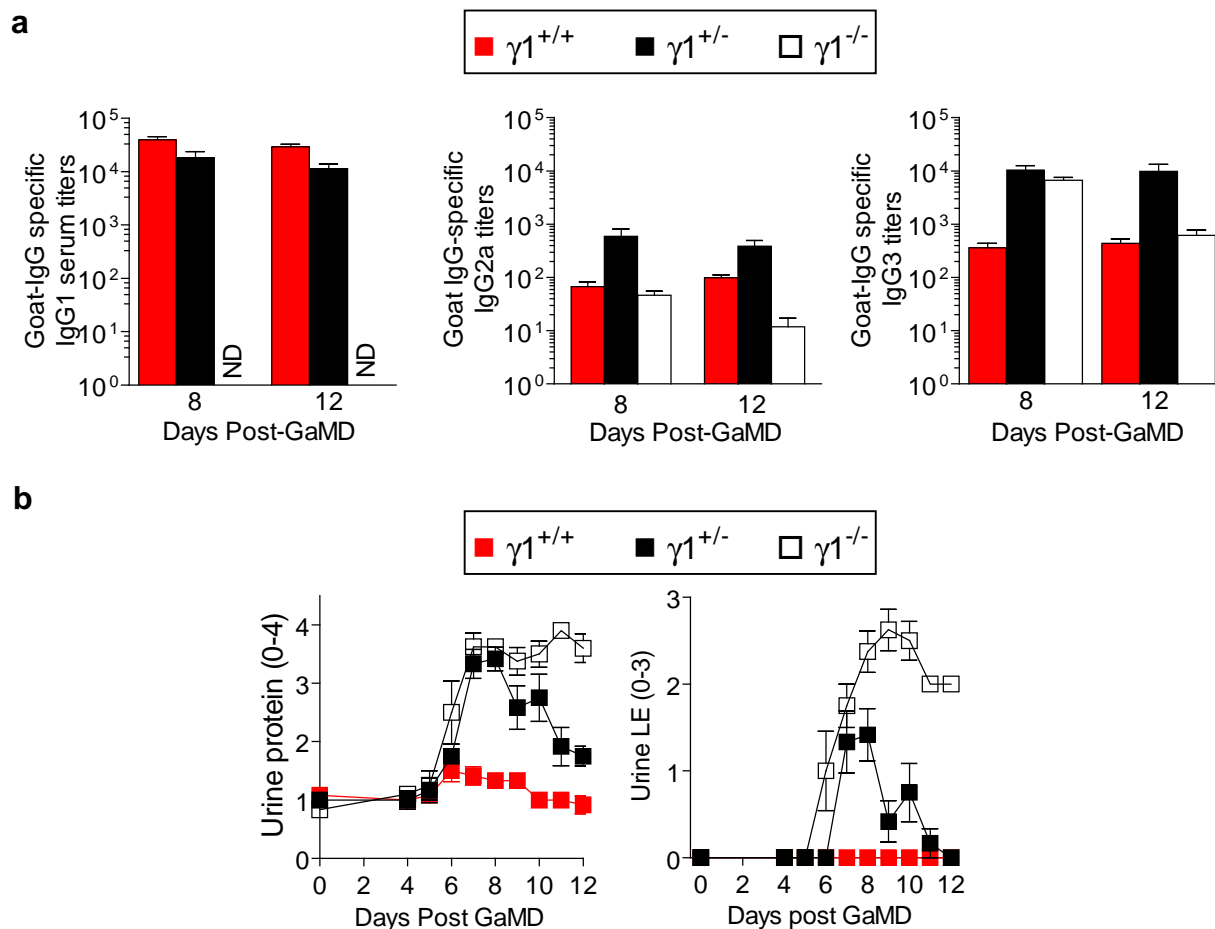
**Extended Data Figure 6 | Severe renal disease develops in GaMD-immunized  $\gamma 1^{-}$ /J-chain $^{-}$  mice.** **a–c**, BALB/c  $\gamma 1^{-}$  (12 mice), J-chain $^{-}$  (9 mice) and  $\gamma 1^{-}$ /J-chain $^{-}$  (12 mice) mice were injected s.c. with GaMD. **a**, Urinalysis

was performed on the indicated days. **b**, BUN levels on day 0 and 11. The difference between  $\gamma 1^{-}$  and  $\gamma 1^{-} \times$  J-chain $^{-}$  mice was not consistently observed. **c**, Survival of GaMD-immunized mice. \* $P < 0.05$ .



**Extended Data Figure 7 | IgG1 inhibits IgG3-induced cryoglobulin kidney disease independent of complement and Fc $\gamma$ RIIB and better than IgG2a and IgG2b.** **a**, Wild-type (WT) mice (4 per group) were injected i.v. with 4 mg of mouse IgG1, IgG2a, IgG2b or IgG3 anti-TNP monoclonal antibody and s.c. with 100  $\mu$ l of TNP-goat serum on days 0 and 1. Urine leukocyte esterase and blood was measured before injections and on day 1 and day 2. **b**, Urine leukocyte esterase (LE) and blood for BALB/c wild-type and C3<sup>-/-</sup> mice (4 per group) injected i.v. with 4 mg of IgG3 anti-TNP monoclonal antibody and s.c. with 400  $\mu$ l of TNP-goat serum on day 0 and day 1. **c**, Wild-type and Fc $\gamma$ RIIB-deficient (Fc $\gamma$ RIIB<sup>-/-</sup>) mice (4 per group) were injected s.c. with 100  $\mu$ l of TNP-goat serum and i.v. with 4 mg of IgG3 anti-TNP plus or minus 5 mg of IgG1 anti-TNP on day 0 and day 1. Urinalysis on days 0, 1 and 2. **d**, BALB/c mice were injected i.v. with 4 mg of IgG3 anti-TNP and s.c. with 1.4 mg of TNP-BSA on days 0 and 1. Some mice were also injected with 0.625, 1.25, 2.5 or

5 mg of switch variants of IgG1, IgG2a or IgG2b anti-TNP monoclonal antibodies on days 0 and 1. Urine protein was determined on day 0 (data not shown), day 1 (top) and day 2 (bottom). Results are pooled from a total of seven experiments. Group size was as follows. IgG3 alone: 19 mice; 0.625 mg of IgG1, IgG2a or IgG2b: 4 mice; 1.25 mg of IgG1, IgG2a or IgG2b: 8 mice; 2.5 mg of IgG1, IgG2a or IgG2b: 6 mice; 5 mg of IgG1, IgG2a or IgG2b: 8 or 9 mice. The significance of differences between treatment groups was determined as described in the legend to Fig. 4f. \* $P$  < 0.05 as compared to IgG3 alone; † $P$  < 0.05 as compared to IgG2b plus IgG3; **e**, Binding of the Ig isotype switch variants to ELISA wells coated with TNP-BSA, reported as percentage of maximal binding. **f**, Binding of the Ig isotype switch variants to ELISA wells coated with IgG3 anti-TNP monoclonal antibody. **g**, Binding of IgG3 and the Ig isotype switch variants to ELISA wells coated with themselves.

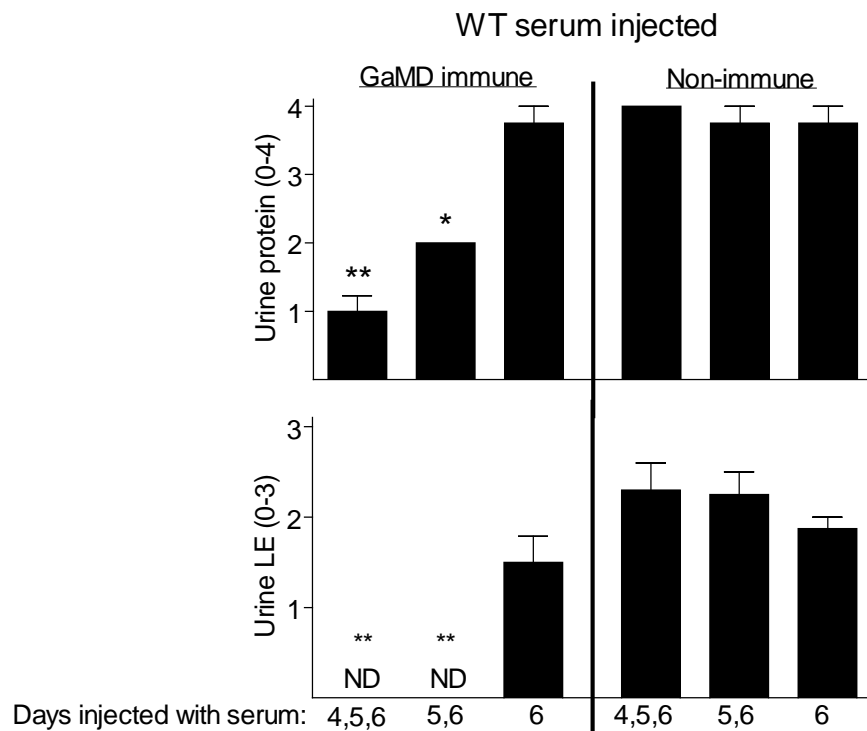


**Extended Data Figure 8 | GaMD-immunized  $\gamma 1^{+/-}$  mice generate large IgG3 responses but develop mild renal disease.** **a, b,** BALB/c mice homozygous ( $\gamma 1^{+/+}$ ), heterozygous ( $\gamma 1^{+/-}$ ) and null ( $\gamma 1^{-/-}$ ) for a functional  $\gamma 1$  allele (6 per group) were injected s.c. with GaMD. **a,** Sera were titred for

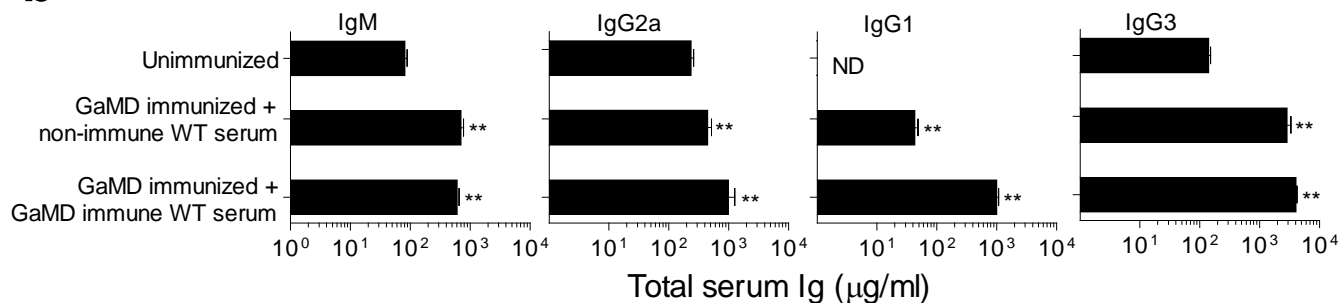
goat IgG-specific IgG1, IgG2a and IgG3 0, 8 and 12 days later. Day 0 titres were zero for all Ig isotypes (data not shown). **b,** Urine samples from the same mice were assayed for protein and leukocyte esterase. ND, none detected.



**a**



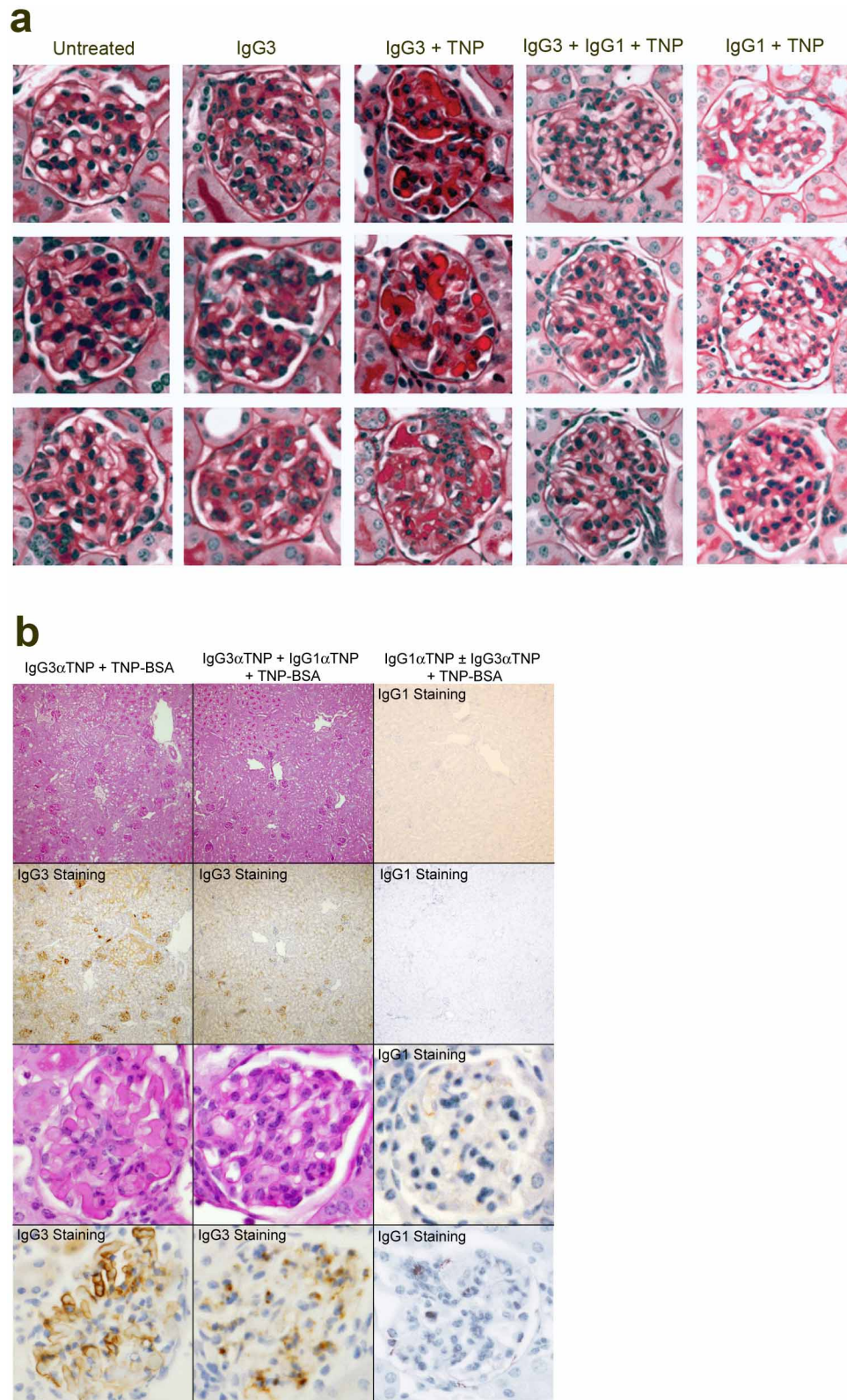
**b**



**Extended Data Figure 9 | GaMD immune serum from wild-type mice inhibits GaMD-induced renal disease without decreasing other isotypes if injected into GaMD-immunized  $\gamma 1^-$  mice by 5 days after immunization.**

**a**, BALB/c  $\gamma 1^-$  mice (4 or 8 per group) were injected s.c. with GaMD on day 0 and i.p. with 0.5 ml of pooled serum from GaMD-immunized wild-type mice (GaMD immune WT serum) or unimmunized wild-type mice (non-immune serum), starting 4, 5 or 6 days after GaMD immunization. Day 7 urine samples were analysed. LE, leukocyte esterase; ND, none detected. **b**, BALB/c

$\gamma 1^-$  mice (4 or 8 per group) were injected s.c. with GaMD on day 0 and i.p. with 0.5 ml of pooled serum from GaMD-immunized wild-type mice (GaMD immune WT serum) or unimmunized wild-type mice (non-immune serum), 5, 6 and 7 days after GaMD immunization. Sera were assayed for total IgG1, IgG2a, IgM and IgG3 on day 0 (unimmunized) and 8 days after GaMD immunization. ND, none detected. \* $P < 0.05$ , \*\* $P < 0.005$  (both as compared to day 6 only in **a** and unimmunized in **b**).



**Extended Data Figure 10 | Antigen-specific IgG1 can prevent IgG3 immune complex glomerular deposition.** **a, b**, BALB/c wild-type (WT) mice were injected i.v. with mouse IgG1 and/or IgG3 anti-TNP monoclonal antibody with or without s.c. injection of TNP-BSA on days 0 and 1. **a**, Kidneys were stained

with PAS on day 2. Representative micrographs from 3 mice per group are shown. **b**, Kidney serial sections were stained with PAS or for IgG3 or IgG1 (brown pigment). Representative micrographs from 4 mice per group are shown. Original magnification,  $\times 200$  for **a, b**.

# Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution

Florian K. M. Schur<sup>1,2</sup>, Wim J. H. Hagen<sup>1</sup>, Michaela Rumlová<sup>3,4</sup>, Tomáš Ruml<sup>5</sup>, Barbara Müller<sup>2,6</sup>, Hans-Georg Kräusslich<sup>2,6</sup> & John A. G. Briggs<sup>1,2</sup>

**Human immunodeficiency virus type 1 (HIV-1) assembly proceeds in two stages. First, the 55 kilodalton viral Gag polyprotein assembles into a hexameric protein lattice at the plasma membrane of the infected cell, inducing budding and release of an immature particle. Second, Gag is cleaved by the viral protease, leading to internal rearrangement of the virus into the mature, infectious form<sup>1</sup>. Immature and mature HIV-1 particles are heterogeneous in size and morphology, preventing high-resolution analysis of their protein arrangement *in situ* by conventional structural biology methods. Here we apply cryo-electron tomography and sub-tomogram averaging methods to resolve the structure of the capsid lattice within intact immature HIV-1 particles at subnanometre resolution, allowing unambiguous positioning of all  $\alpha$ -helices. The resulting model reveals tertiary and quaternary structural interactions that mediate HIV-1 assembly. Strikingly, these interactions differ from those predicted by the current model based on *in vitro*-assembled arrays of Gag-derived proteins from Mason–Pfizer monkey virus<sup>2</sup>. To validate this difference, we solve the structure of the capsid lattice within intact immature Mason–Pfizer monkey virus particles. Comparison with the immature HIV-1 structure reveals that retroviral capsid proteins, while having conserved tertiary structures, adopt different quaternary arrangements during virus assembly. The approach demonstrated here should be applicable to determine structures of other proteins at subnanometre resolution within heterogeneous environments.**

HIV-1 Gag consists of separately folded domains connected by flexible linkers. Its amino (N)-terminal matrix (MA) domain directs Gag to the plasma membrane, and the subsequent two capsid (CA) domains mediate protein–protein interactions in the mature and immature particle; they are followed by the RNA-binding nucleocapsid (NC) domain. Spacer-peptides SP1 and SP2 are positioned between CA and NC, and between NC and the carboxy (C)-terminal p6 domain, respectively<sup>3</sup>. The immature protein shell constitutes a hexameric Gag array bound to the viral membrane, forming a truncated sphere with irregular defects<sup>4</sup>. Proteolytic cleavage at five positions in Gag during maturation leads to a dramatic structural rearrangement, with cleaved CA forming the characteristic conical HIV-1 core encapsulating the now condensed NC–RNA complex.

High-resolution structures of all isolated Gag domains are available<sup>5</sup>, and a structural model of the mature HIV-1 capsid core has recently been generated<sup>6</sup>. However, the flexible nature of the uncleaved polyprotein has made structural studies of the immature Gag conformation more challenging. We recently described the structure of an *in vitro*-assembled immature-like tubular array of a Mason–Pfizer monkey virus (M–PMV) Gag-derived protein at  $\sim 8$  Å resolution<sup>2</sup>, which currently represents the most detailed model for an immature Gag lattice. For immature HIV-1, a partial model has been generated based on *in vitro*-assembled tubular arrays of Gag-derived protein in which a mutation stabilizes the C-terminal CA domain (CA-CTD) in an immature-like

form, while the N-terminal CA domain (CA-NTD) adopts a presumably non-physiological form<sup>7</sup>.

Subtomogram averaging has been used to solve low-resolution structures of biological molecules within pleiotropic native environments<sup>8</sup>, including structures of the immature HIV-1 Gag lattice within virus particles at  $\sim 20$  Å resolution<sup>4,9</sup>. Higher-resolution structures have not yet been obtained for any component of such heterogeneous viruses *in situ*.

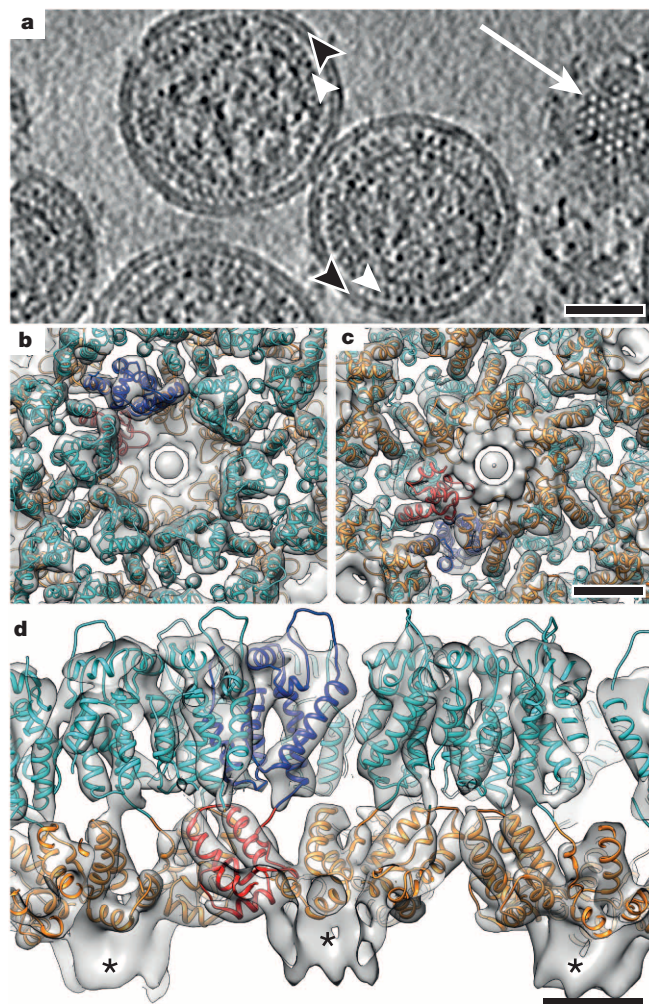
Recently, we showed that optimized subtomogram averaging methods can recover structural data at below 10 Å resolution<sup>10</sup>. Their application to immature HIV-1 might enable determination of a subnanometre-resolution structure of Gag within intact virus. Therefore, we purified immature HIV-1 particles treated with the protease inhibitor amprenavir (APV) (Extended Data Fig. 1a) and prepared them for cryo-electron tomography. We collected and reconstructed 70 cryo-electron tomograms as described<sup>10</sup> and segregated them into two data sets for independent processing. Reconstructed tomograms contained 224 immature HIV-1 particles with variable diameters and the characteristic striated density corresponding to the Gag lattice<sup>11</sup> (Fig. 1a, Extended Data Fig. 2a and Extended Data Table 1). Subtomograms extracted along the capsid layers were iteratively aligned and averaged in three dimensions. The aligned positions of subtomograms within individual virus particles indicated a continuous but incomplete  $\sim 8$  nm spaced hexameric lattice as previously described<sup>4</sup> (Extended Data Fig. 2a). The CA layer within the final average of all data was resolved at 8.8 Å (Extended Data Fig. 2b), showing well-resolved rod-shaped densities for  $\alpha$ -helices (Fig. 1b–d and Extended Data Fig. 2d–f). The general domain arrangement seen in the APV-inhibited HIV-1 was confirmed in an 11 Å resolution structure of HIV-1 with an inactivating protease mutation (Extended Data Figs 1b and 2d–f).

The final structure shows hexamerically arranged densities in the CA and SP1 regions, similar to those previously described at low resolution<sup>4,9</sup> (Fig. 1b–d). No well-ordered densities were observed for MA or NC (Extended Data Fig. 2c), indicating that they do not adopt a defined spatial arrangement relative to CA, consistent with the flexibility of the inter-domain linkers. In the region presumably corresponding to SP1, six rod-like densities were observed, arranged in a tube around a hollow centre at the sixfold symmetry axis (Fig. 1d and Supplementary Video 1), consistent with previous models suggesting this region assembles a six-helix bundle<sup>7,9,12</sup>.

We performed rigid-body fitting of high-resolution structures for CA-NTD (Protein Data Bank (PDB) 1L6N<sup>13</sup>) and CA-CTD (PDB 3DS2 (ref. 14)) into the outer and inner layers of the CA density, respectively, and refined the fit using molecular dynamics flexible fitting<sup>15</sup> (Extended Data Fig. 3a, b). This provided a model showing the tertiary and quaternary interactions in the CA layer of the immature virus (Figs 1b–d, 2 and Supplementary Video 1). For both domains a single unambiguous fit was identified accommodating all 11  $\alpha$ -helices of CA within appropriate

<sup>1</sup>Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. <sup>2</sup>Molecular Medicine Partnership Unit, European Molecular Biology Laboratory/Universitätsklinikum Heidelberg, Heidelberg, Germany. <sup>3</sup>Institute of Organic Chemistry and Biochemistry (IOCB), Academy of Sciences of the Czech Republic, v.v.i., IOCB & Gilead Research Center, Flemingovo nám. 2, 166 10 Prague, Czech Republic. <sup>4</sup>Department of Biotechnology, Institute of Chemical Technology, Prague, Technická 5, 166 28, Prague, Czech Republic. <sup>5</sup>Department of Biochemistry and Microbiology, Institute of Chemical Technology, Prague, Technická 5, 166 28, Prague, Czech Republic. <sup>6</sup>Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany.



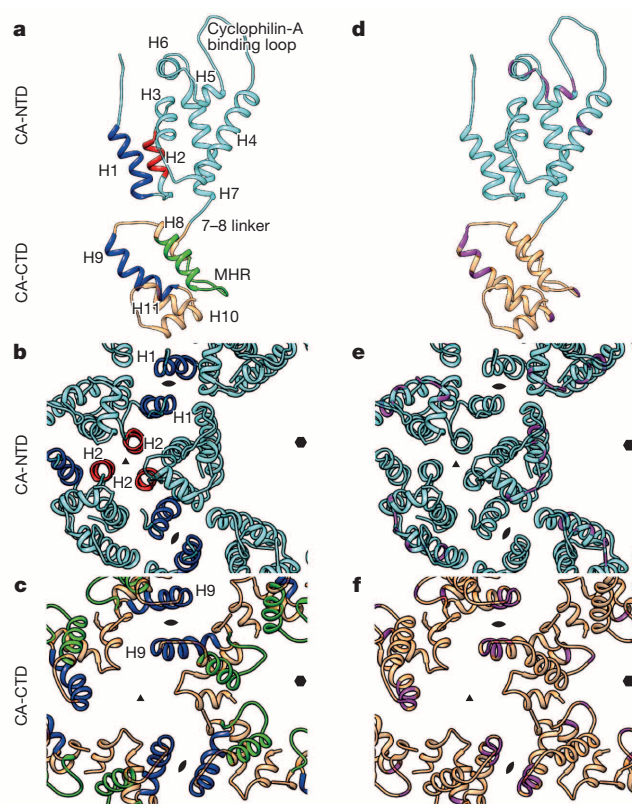


**Figure 1 | Structure of the CA-SP1 lattice within immature HIV-1.**

**a**, Computational slice through a Gaussian-filtered tomogram containing immature HIV-1 particles treated with APV. White arrowheads indicate the immature CA layer; black arrowheads indicate the membrane; white arrow marks a grazing slice through the CA layer illustrating the hexagonal lattice. Scale bar, 50 nm. **b–d**, Isosurface representation of the final structure showing the immature CA-SP1 lattice viewed from outside the virus (**b**), from inside (**c**) and in an orthogonal view (**d**). Isosurface threshold value is  $2\sigma$  away from the mean. High-resolution structures for CA-NTD (cyan) and CA-CTD (orange) have been fitted into the density. An individual capsid monomer is coloured blue/red. Unfilled densities marked with asterisks correspond to the SP1 region. Scale bar, 25 Å. See also Supplementary Video 1.

densities in the electron microscopy map. At lower isosurface thresholds, densities corresponding to the peptide chain linking helix 7 in CA-NTD with helix 8 in CA-CTD, as well as the cyclophilin-A binding loop, were observed (Extended Data Fig. 3b–d and Supplementary Video 2).

The arrangement of the CA-CTD is similar to those in partly immature-like *in vitro*-assembled tubular arrays of HIV-1 (ref. 7) (Extended Data Fig. 4) and tubular arrays of M-PMV Gag<sup>2</sup>. Between hexamers, two CA-CTD domains form a homo-dimeric interface mediated by helix 9 (Fig. 2c). The positions of assembly-critical hydrophobic residues W316 and M317 in helix 9 (refs 16, 17) are consistent with a contribution to the interaction interface (Fig. 2f). Around the hexamer, contacts between adjacent CA-CTD monomers include part of the major homology region (I285–L304)<sup>18</sup>, consistent with a role for this region in stabilizing the hexamer. Other residues, where mutations are known to disrupt viral assembly<sup>16,17</sup>, are close to inter-protein interfaces (Fig. 2d–f). These observations point to the central role of CA-CTD in assembling and stabilizing the immature lattice.

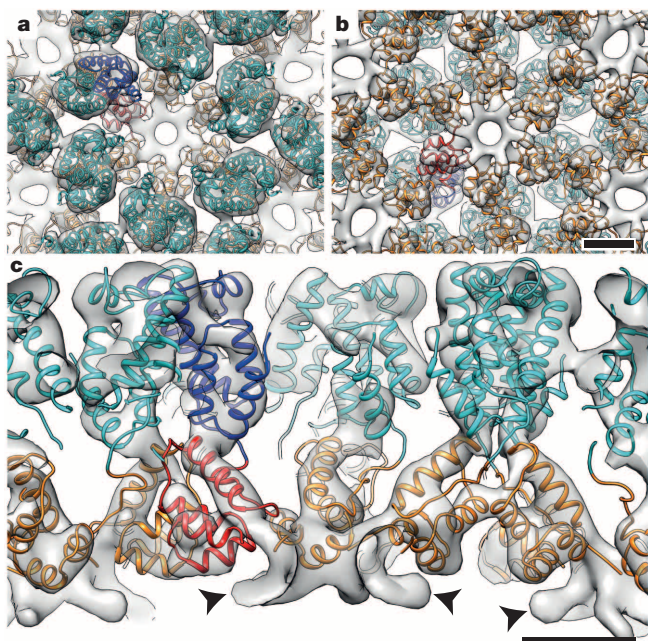


**Figure 2 | Interactions stabilizing the immature HIV-1 capsid lattice.** **a**, An annotated CA monomer, indicating helices contributing to homo-dimeric and trimeric CA-NTD and CA-CTD interactions in blue and red, respectively. Interactions around the hexamer involve residues in the major homology region (green). **b**, The CA-NTD can form intra-hexameric interactions, dimeric inter-hexameric contacts between helices 1 and trimeric inter-hexameric interactions between helices 2. **c**, In the CA-CTD a dimeric interaction between helices 9 is observed. **d–f**, Views as in **a–c**: residues in which mutations disrupt particle assembly are coloured magenta. Sixfold, threefold and twofold symmetry axes are indicated by black hexagons, triangles and ellipses, respectively.

The CA-NTD and CA-CTD abut each other at two positions in the immature lattice (Supplementary Video 2): close to residues S310–Q311 in helix 9 and E161–K162 in helix 1, and close to residues M276–S278 in helix 7 and K302–R305 in helix 8 (Extended Data Fig. 3c, d). Unlike in the mature capsid, there is no large CA-CTD:CA-NTD interface<sup>19,20</sup>. This is consistent with the ability of CA-CTD to assemble an immature-like arrangement independent of the CA-NTD arrangement *in vitro*<sup>7</sup>, and to induce release of virus-like particles from cells in Gag protein mutants lacking CA-NTD<sup>21</sup>.

Virus-like particles lacking CA-NTD have aberrant size distributions<sup>21</sup>, indicating a role for CA-NTD in determining particle morphology. Within the immature virus, each CA-NTD domain can interact with five neighbouring CA-NTD domains (Fig. 2b). Around the hexamer, the top of helix 4 in each CA-NTD interacts with a patch of residues between helices 5 and 6 in the neighbouring CA-NTD (Supplementary Video 2). A homo-dimeric interface is formed between helix 1 from two adjacent hexamers. A homo-trimeric, three-helix bundle interface is formed by helix 2 from neighbouring hexamers. The weaker density of helix 2 compared with other helices within our electron microscopy density suggests it may have some flexibility within the lattice. Together these interfaces comprise an extensive network of interactions. The integrity of the lattice could be maintained by any combination of two of these three interfaces, and mutation in a single interface may therefore not interfere with assembly. This may explain why mutation of CA-NTD residues close to inter-hexamer protein–protein interfaces has a mild or no effect on virus assembly<sup>16,22</sup>.



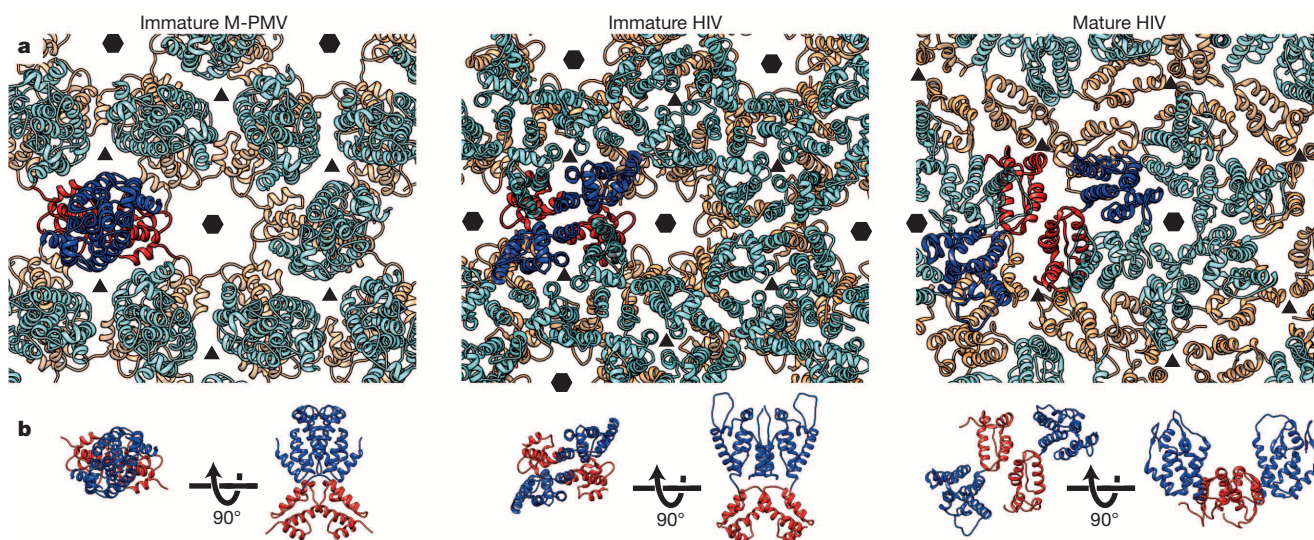


**Figure 3 | Structure of the capsid lattice within immature M-PMV (D26N) particles.** **a–c**, Isosurface representation of the immature M-PMV CA lattice viewed from outside the virus (**a**), from inside (**b**) and in an orthogonal view (**c**). Partly resolved  $\alpha$ -helices are visible, allowing positioning of high-resolution structures for CA-NTD and CA-CTD into the density. Colouring as in Fig. 1. The splayed densities annotated with black arrowheads correspond to the 'spacer-peptide-like' region in M-PMV. Isosurface threshold is  $2.5\sigma$  away from the mean. Scale bar, 25 Å.

Mutation of certain residues in helices 4–6 has been shown to interfere with virus assembly<sup>16</sup>. While some of these residues may contribute to the intra-hexamer interfaces, others face the central hole of the hexamer (Fig. 2e). They may modulate the structural integrity of CA-NTD, represent interfaces required transiently during assembly, or be a binding site for unknown host factors regulating assembly. The arrangement of CA-NTD domains leaves reported binding sites of host cofactors such as cyclophilin-A accessible<sup>23</sup>.

Surprisingly, the arrangement of CA-NTD domains is completely different from the current model for immature retroviruses based on the structure of *in vitro*-assembled tubular arrays of M-PMV Gag<sup>2</sup>. This could be due to a true difference between the immature lattices in HIV-1 and M-PMV, or because the tubular array structure represents an artefact of *in vitro* assembly. We purified intact immature M-PMV containing an inactivating protease mutation (D26N) (Extended Data Figs 1c and 5a). Following the same protocol as for HIV, we obtained a 9.7 Å resolution reconstruction of the CA layer (Extended Data Fig. 5b–f) from which we generated a structural model (Fig. 3). The arrangement of CA domains in intact immature M-PMV matches the arrangement within *in vitro*-assembled M-PMV tubular arrays<sup>2</sup> (Extended Data Fig. 6), and is indeed different from that observed in immature HIV-1 (Figs 1b–d and 4). Thus, although retroviral CA proteins have conserved tertiary structures (Extended Data Fig. 7), they adopt different quaternary arrangements within the immature virus (Extended Data Fig. 8). While HIV-1 CA-NTD forms an extensive network of interactions linking within and between hexamers, the primary interaction of M-PMV CA-NTD is a larger dimeric interface linking two neighbouring hexamers. How these different arrangements have evolved can only be speculated. We suspect their functional relevance relates to the site of assembly: the lentivirus HIV-1 makes use of the plasma membrane as a scaffold, while the betaretrovirus M-PMV assembles within the cytoplasm<sup>24</sup> and may require stronger interactions to link Gag hexamers.

The arrangement of CA in immature HIV-1 is very different from that suggested for the mature HIV-1 capsid core<sup>6,19,25</sup> (Extended Data Fig. 8). In both immature and mature states, CA-CTD dimerization is mediated by helix 9, but with a different relative orientation of the two monomers (Fig. 4 and Extended Data Fig. 4d). This and the dramatic rearrangement of the CA-NTD from immature to mature particles would require breaking essentially all interactions of the immature lattice and forming new contacts during maturation (Supplementary Video 3). This could be achieved by two routes: (1) by disassembly of the immature lattice into monomers or dimers linked by the helix 9 interface; (2) by disassembly of the immature lattice into hexamers or higher-order oligomers. Route 2 would imply a two-step 'molecular dance'. First, the hexameric interactions in the CA-CTD-SP1 region (which can form even when CA-NTD is not properly arranged<sup>7</sup>) could stabilize the hexamer while cleavage between MA and CA induces the CA-NTD to move around the flexible 7–8 linker into its mature hexameric arrangement.



**Figure 4 | Comparison of the CA arrangement in immature and mature retroviruses.** **a**, The arrangement of the CA lattices in immature M-PMV and immature and mature HIV particles. CA-NTD and CA-CTD domains are coloured in cyan/blue and orange/red, respectively. See also Supplementary Video 3. Symmetry axes are marked as in Fig. 2. **b**, The structures of CA dimers

from the lattices in **a** are shown in top and orthogonal views. CA-CTD interactions are always mediated by helices 9, although the dimer rotates around this interaction in the transition from the immature to the mature form (see Extended Data Fig. 4). CA-NTD interactions are in each case entirely different. See also Extended Data Fig. 8.

Second, these now mature CA-NTD interactions would stabilize the hexamer while cleavage between CA and SP1 induces maturation of CA-CTD.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 10 May; accepted 3 September 2014.**

**Published online 2 November 2014.**

1. Sundquist, W. I. & Krausslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb. Perspec. Med.* **2**, a006924 (2012).
2. Bharat, T. A. *et al.* Structure of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy. *Nature* **487**, 385–389 (2012).
3. Bell, N. M. & Lever, A. M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol.* **21**, 136–144 (2013).
4. Briggs, J. A. *et al.* Structure and assembly of immature HIV. *Proc. Natl Acad. Sci. USA* **106**, 11090–11095 (2009).
5. Briggs, J. A. & Krausslich, H. G. The molecular architecture of HIV. *J. Mol. Biol.* **410**, 491–500 (2011).
6. Zhao, G. *et al.* Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646 (2013).
7. Bharat, T. A. *et al.* Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *Proc. Natl Acad. Sci. USA* **111**, 8233–8238 (2014).
8. Briggs, J. A. Structural biology *in situ* — the potential of subtomogram averaging. *Curr. Opin. Struct. Biol.* **23**, 261–267 (2013).
9. Wright, E. R. *et al.* Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J.* **26**, 2218–2226 (2007).
10. Schur, F. K., Hagen, W. J., de Marco, A. & Briggs, J. A. Determination of protein structure at 8.5 Å resolution using cryo-electron tomography and sub-tomogram averaging. *J. Struct. Biol.* **184**, 394–400 (2013).
11. Fuller, S. D., Wilk, T., Gowen, B. E., Kräusslich, H.-G. & Vogt, V. M. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.* **7**, 729–738 (1997).
12. Accola, M. A., Höglund, S. & Göttinger, H. G. A. Putative  $\alpha$ -helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 gag precursor is crucial for viral particle assembly. *J. Virol.* **72**, 2072–2078 (1998).
13. Tang, C., Ndassa, Y. & Summers, M. F. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nature Struct. Biol.* **9**, 537–543 (2002).
14. Bartonova, V. *et al.* Residues in the HIV-1 capsid assembly inhibitor binding site are essential for maintaining the assembly-competent quaternary structure of the capsid protein. *J. Biol. Chem.* **283**, 32024–32033 (2008).
15. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B. & Schulten, K. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174–180 (2009).
16. von Schwedler, U. K., Stray, K. M., Garrus, J. E. & Sundquist, W. I. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J. Virol.* **77**, 5439–5450 (2003).
17. Chu, H.-H., Chang, Y.-F. & Wang, C.-T. Mutations in the  $\alpha$ -helix directly C-terminal to the major homology region of human immunodeficiency virus type 1 capsid protein disrupt gag multimerization and markedly impair virus particle production. *J. Biomed. Sci.* **13**, 645–656 (2006).
18. Mammano, F., Ohagen, A., Höglund, S. & Göttinger, H. G. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.* **68**, 4927–4936 (1994).
19. Pornillos, O. *et al.* X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282–1292 (2009).
20. Monroe, E. B., Kang, S., Kyere, S. K., Li, R. & Prevelige, P. E., Jr. Hydrogen/deuterium exchange analysis of HIV-1 capsid assembly and maturation. *Structure* **18**, 1483–1491 (2010).
21. Borsetti, A., Ohagen, A. & Göttinger, H. G. The C-terminal half of the human immunodeficiency virus type 1 gag precursor is sufficient for efficient particle assembly. *J. Virol.* **72**, 9313–9317 (1998).
22. von Schwedler, U. K. *et al.* Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555–1568 (1998).
23. Franke, E. K., Yuan, H. E. H. & Luban, J. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **372**, 359–362 (1994).
24. Chopra, H. C. & Mason, M. M. A new virus in a spontaneous mammary tumor of a rhesus monkey. *Cancer Res.* **30**, 2081–2086 (1970).
25. Pornillos, O., Ganser-Pornillos, B. K. & Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **469**, 424–427 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This study was supported by Deutsche Forschungsgemeinschaft grants BR 3635/2-1 to J.A.G.B., KR 906/7-1 to H.-G.K. and by Grant Agency of the Czech Republic 14-15326S to M.R. The Briggs laboratory acknowledges financial support from the European Molecular Biology Laboratory and from the Chica und Heinz Schaller Stiftung. We thank B. Glass, M. Anders and S. Mattei for preparation of samples, and R. Hadravova, K. H. Bui, F. Thommen, M. Schorb, S. Dodonova, S. Glatt, P. Ulbrich and T. Bharat for technical support and/or discussion. This study was technically supported by the European Molecular Biology Laboratory IT services unit.

**Author Contributions** F.K.M.S., M.R., T.R., B.M., H.-G.K. and J.A.G.B. designed and interpreted experiments. F.K.M.S. and W.J.H.H. collected data, F.K.M.S. performed image processing, and F.K.M.S. and J.A.G.B. analysed data. F.K.M.S. and J.A.G.B. wrote the manuscript with support from all authors.

**Author Information** Cryo-electron microscopy structures and a representative tomogram have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-2706, EMD-2707 and EMD-2708, and the fitted HIV atomic model in the PDB under accession number 4USN. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.G.B. ([john.briggs@embl.de](mailto:john.briggs@embl.de)).



## METHODS

**Virus particle production and purification.** HEK293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (FCS; Biobrom), penicillin ( $100 \text{ IU ml}^{-1}$ ), streptomycin ( $100 \mu\text{g ml}^{-1}$ ) and 4 mM glutamine. For the protease-inhibitor-treated sample, cells were transfected with proviral plasmid pNL4-3 (ref. 26) using polyethyleneimine following standard procedures. A final concentration of  $5 \mu\text{M}$  amprenavir<sup>27</sup> (obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health) was added to the tissue culture medium at the time of transfection. For the protease mutant HIV and M-PMV sample, cells were transfected with plasmids pNL4-3(PR-) (in case of HIV D25A) or pSHRM15 (for M-PMV D26N). At 44 hours post-transfection, tissue culture supernatant was harvested and filtered through  $0.45 \mu\text{m}$  nitrocellulose filters. Virus was enriched by ultracentrifugation through a 20% (w/w) sucrose cushion and further purified by centrifugation through an iodixanol density gradient<sup>28</sup>. Concentrated virus samples were re-suspended in 25 mM MES pH 6.0, 150 mM NaCl or phosphate buffered saline (PBS), and treated with 1.2% paraformaldehyde for 1 h on ice and stored in aliquots at  $-80^\circ\text{C}$ . Purities of samples and Gag maturation state were verified by SDS–polyacrylamide gel electrophoresis followed by silver staining and immunoblotting (Extended Data Fig. 1).

**Cryo-electron microscopy and image processing.** Degassed C-Flat 2/2-2C or 2/2-3C grids were glow discharged for 30 s at 20 mA. Virus solution was diluted with PBS containing 10 nm colloidal gold and 2  $\mu\text{l}$  of this mixture was then applied to grids and plunge frozen in liquid ethane using an FEI Vitrobot Mark 2. The grids were stored under liquid nitrogen conditions until imaging.

Data acquisition and image processing for immature HIV treated with APV and the immature M-PMV protease mutant (D26N) was performed as previously described<sup>10</sup> (Extended Data Table 1). In brief, tilt series were imaged on a FEI Titan Krios electron microscope operated at 200 keV, with a GIF2002 post-column energy filter (using a slit width of 20 eV) and a  $2\text{K} \times 2\text{K}$  Gatan Multiscan 795 CCD (charge-coupled device) camera. For navigation and search purposes, low-magnification montages were acquired using SerialEM<sup>29</sup>, and tilt series were acquired at appropriate positions using FEI tomography software version 4 in automated batch mode. The nominal magnification was 42,000, giving a calibrated pixel size of  $2.025 \text{ \AA}$ . The tilt range was from  $-45^\circ$  to  $+60^\circ$  in  $3^\circ$  steps, collecting first from  $0^\circ$  to  $-45^\circ$  and then from  $3^\circ$  to  $60^\circ$ . Tilt series were collected at a range of nominal defoci between  $-1.2$  and  $-4.5 \mu\text{m}$ . The total dose applied to each tilt series was approximately  $40 \text{ e}^- \text{ \AA}^{-2}$ . Tomograms were reconstructed using the IMOD software suite<sup>30</sup>.

Image acquisition for protease defective HIV (D25A) was performed on an FEI Titan Krios electron microscope operated at 300 keV equipped with a Falcon II direct electron detector. Acquisition parameters were the same as described above, but using defoci ranging from  $-2.0$  to  $-5.5 \mu\text{m}$  and having a calibrated pixel size of  $2.281 \text{ \AA}$  per pixel. The total dose was approximately  $60 \text{ e}^- \text{ \AA}^{-2}$ . The tilt range was either from  $-45^\circ$  to  $+60^\circ$  in  $3^\circ$  steps or from  $-42^\circ$  to  $+42^\circ$  in  $3^\circ$  steps.

Image processing was performed identically for all samples. Before image processing, tomograms were split into two half-data sets with each set containing roughly the same number of virus particles and the same distribution of defoci, to obtain two independent reconstructions. Subtomogram averaging calculations were performed using Matlab scripts derived from the AV3 (ref. 31) and TOM<sup>32</sup> packages. Generation of masks and alignment of the two final references before Fourier shell correlation (FSC) measurement was done using the Dynamo software package<sup>33</sup>. Visualization of tomograms was performed in either IMOD<sup>30</sup> or Amira (FEI Visualization Sciences Group) using the electron microscopy toolbox<sup>34</sup>.

Initial processing was performed on  $4 \times$  binned, non-contrast transfer function (CTF) corrected data. For each half-data set, one tomogram was chosen to obtain an initial structure. Cubic subtomograms with a size of  $(340 \text{ \AA})^3$  were extracted from the surface of a sphere with its centre being the centre of the virus and the radius corresponding to the mean radius at the CA-level. The subtomograms were assigned initial angles based on the geometry of the sphere and then averaged to obtain a smooth reference. Subtomograms were aligned and averaged in an iterative manner against this reference in six dimensions to obtain an initial structure<sup>31</sup>. Subsequently all subtomograms within each half-data set were aligned and averaged against their respective independent starting reference for six iterations. After the first two iterations a cross-correlation-based cleaning was performed to remove subtomograms that contained no density corresponding to the Gag layer.

The defocus of each tomogram was measured by fitting theoretical CTF-curves to averaged power spectra from  $512 \text{ pixel} \times 512 \text{ pixel}$  tiles generated from all images in a tilt series using Matlab. CTF correction used the program 'ctf phase flip' implemented in IMOD<sup>35</sup>. Subvolumes with a size of  $(284 \text{ \AA})^3$  were extracted from unbinned, CTF-corrected tomograms at the positions determined in the  $4 \times$  binned alignments, and subjected to four further alignment iterations. The sixfold symmetry inherent in the structure was applied at each step. The two final references were aligned, multiplied with a Gaussian filtered mask and the FSC was calculated.

For the APV-treated HIV sample, comparison of the averages of all subtomograms after alignment ( $174,360$  and  $168,144$  asymmetric units in each of the half-data sets, respectively) by FSC indicated a resolution of  $8.9 \text{ \AA}$ . After removing subtomograms with lower cross-correlation values, final averages were generated from  $94,986$  and  $99,744$  asymmetric units in each of the half-data sets, giving a slightly improved resolution of  $8.8 \text{ \AA}$ , but not perceptibly changing the structure. Subsequently, we averaged the two half-data set maps and sharpened them with an empirically determined negative B-factor of  $-1,200 \text{ \AA}^2$ , while filtering to the resolution determined at the  $0.143$  FSC threshold<sup>36</sup>.

For the M-PMV protease mutant (D26N) sample, 39 tomograms were reconstructed from which 84 viruses were then extracted. The processing was performed as for the protease-inhibitor treated HIV sample. A total of  $77,520$  asymmetric units from both half-data sets were averaged to generate the final volume and the resolution was determined by FSC to be  $9.7 \text{ \AA}$  at the  $0.143$  criterion. Finally the structure was sharpened applying a negative B-factor of  $-1,000 \text{ \AA}^2$ , while filtering to the resolution determined at the  $0.143$  FSC threshold.

For a summary of image-processing statistics and information on the protease defective HIV (D25A) sample, see Extended Data Table 1.

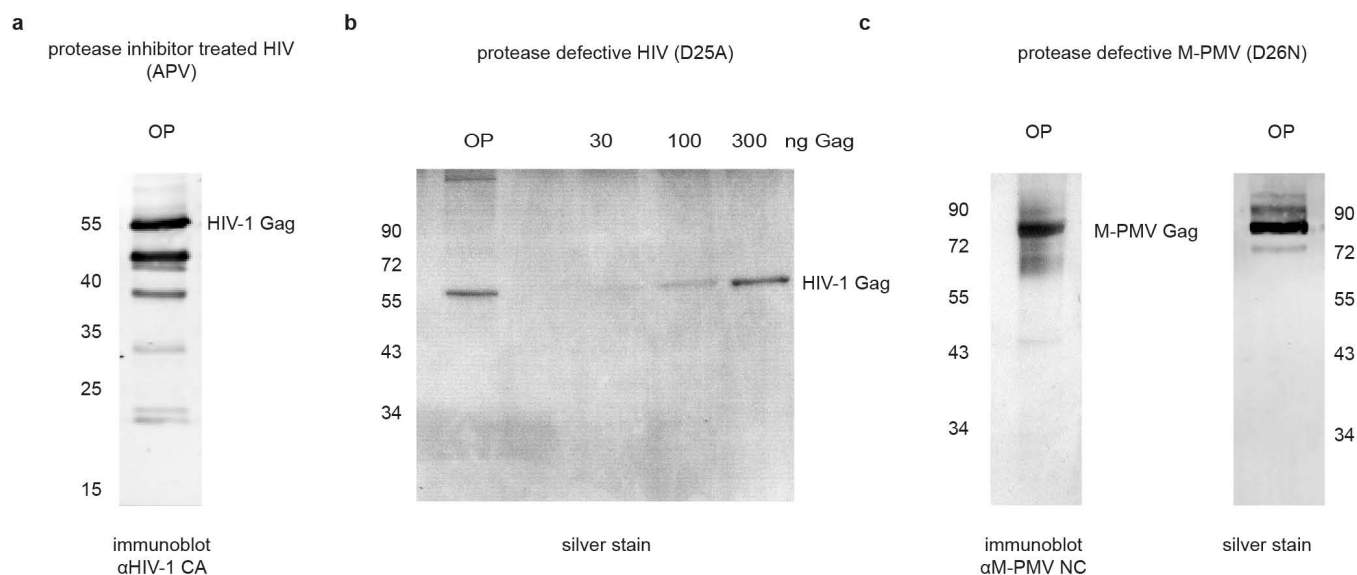
**Structure fitting and analysis.** Visualization and rigid body docking was performed in UCSF Chimera<sup>37</sup>. For flexible fitting into the protease-inhibitor treated HIV map, the nuclear magnetic resonance structure for CA-NTD (PDB 1L6N, chain 1) and the crystal structure for CA-CTD (PDB 3DS2, one monomer) were used. Initially, both models were rigid-body docked into the electron microscopy density to establish their correct relative position using the 'Fit in map' option and were then joined using the 'Adjust bond' option in Chimera. We truncated the joined model from residues 142 to 351 and added PDB 1U57 (ref. 38) (chain 3) (residues GVGPG HKARVLAEMSQVT) to generate a preliminary model of the SP1 region. Finally, 30 copies of the joined monomers were rigid-body docked into the electron microscopy map. The outer 12 capsid monomers and all residues corresponding to SP1 were included in the simulations to prevent excessive movement of residues into empty densities but were removed after the simulation and not interpreted. All flexible fitting simulations were performed using molecular dynamics flexible fitting<sup>15</sup> and NAMD version 2.9 (ref. 39). The simulations were performed until the backbone atoms stopped moving and the fit stabilized. The flexible fitting resulted in only minor changes (Extended Data Fig. 3a, b), highlighting the strong initial correspondence between our map and published models of the individual domains.

To generate the structural model for M-PMV, the PDB for CA-NTD 2KGF<sup>40</sup> (chain 1, residues 317–439) was combined with a homology model of M-PMV CA-CTD (residues 440–516) based on PDB 3H4E<sup>19</sup> (chain L) and fitted into the volume as a rigid body. The mature HIV-1 structure presented for comparison is PDB 3J34 (ref. 6).

26. Bohne, J. & Kräusslich, H.-G. Mutation of the major 5' splice site renders a CMV-driven HIV-1 proviral clone Tat-dependent: connections between transcription and splicing. *FEBS Lett.* **563**, 113–118 (2004).
27. Fung, H. B., Kirschenbaum, H. L. & Hameed, R. Amprenavir: a new human immunodeficiency virus type 1 protease inhibitor. *Clin. Ther.* **22**, 549–572 (2000).
28. Dettenhofer, M. & Yu, X.-F. Highly purified human immunodeficiency virus type 1 reveals a virtual absence of Vif in virions. *J. Virol.* **73**, 1460–1467 (1999).
29. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
30. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
31. Forster, F., Medalia, O., Zauberman, N., Baumeister, W. & Fass, D. Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc. Natl Acad. Sci. USA* **102**, 4729–4734 (2005).
32. Nickell, S. et al. TOM software toolbox: acquisition and analysis for electron tomography. *J. Struct. Biol.* **149**, 227–234 (2005).
33. Castano-Diez, D., Kudryashev, M., Arheit, M. & Stahlberg, H. Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *J. Struct. Biol.* **178**, 139–151 (2012).
34. Pruggnaller, S., Mayr, M. & Frangakis, A. S. A visualization and segmentation toolbox for electron microscopy. *J. Struct. Biol.* **164**, 161–165 (2008).
35. Xiong, Q., Morpheus, M. K., Schwartz, C. L., Hoenger, A. H. & Mastronarde, D. N. CTF determination and correction for low dose tomographic tilt series. *J. Struct. Biol.* **168**, 378–387 (2009).
36. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
37. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
38. Morellet, N., Druillennec, S., Lenoir, C., Bouaziz, S. & Roques, B. P. Helical structure determined by NMR of the HIV-1 (345–392)Gag sequence, surrounding p2: implications for particle assembly and RNA packaging. *Protein Sci.* **14**, 375–386 (2005).
39. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
40. Macek, P. et al. NMR structure of the N-terminal domain of capsid protein from the Mason-Pfizer monkey virus. *J. Mol. Biol.* **392**, 100–114 (2009).

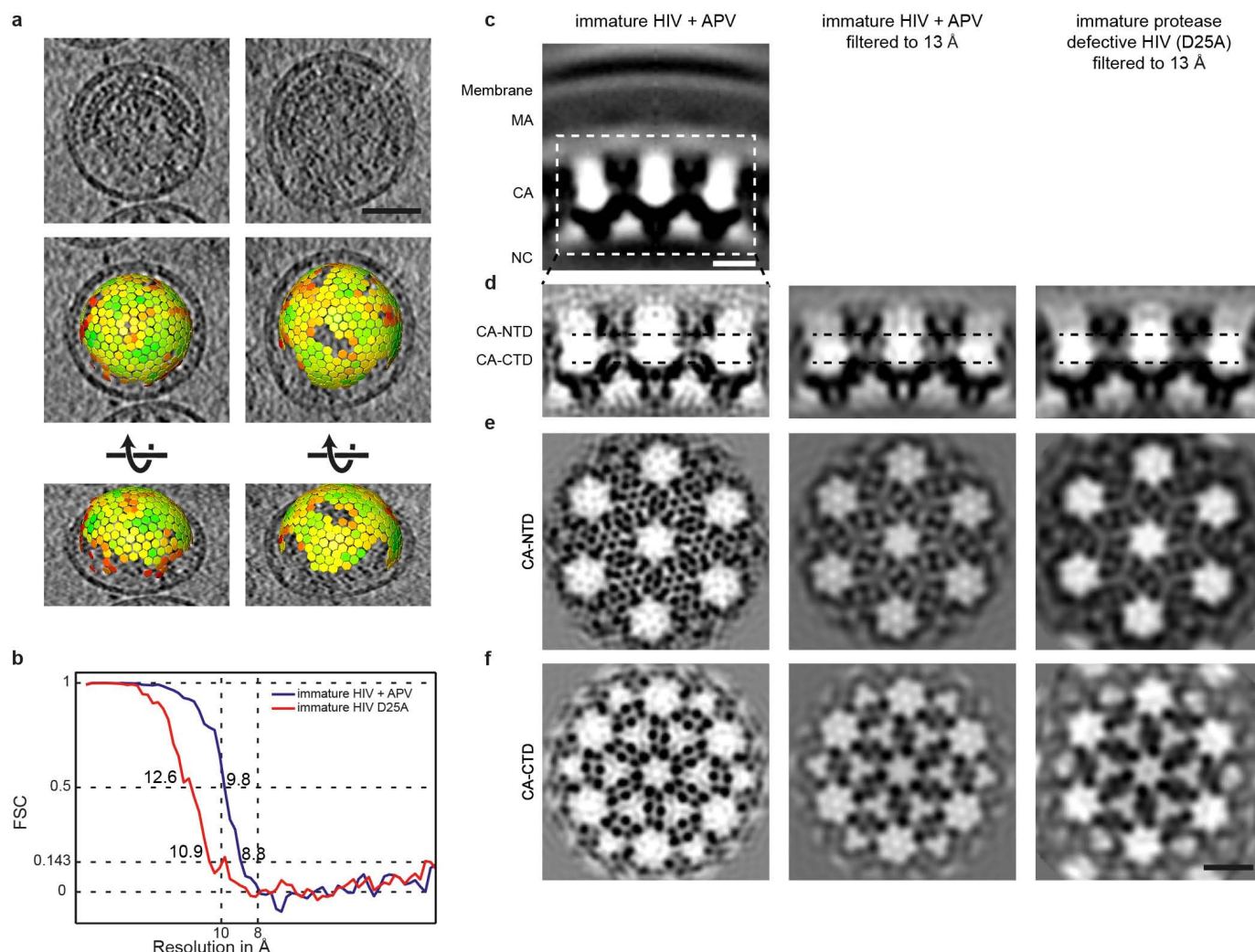
41. de Marco, A. *et al.* Role of the SP2 domain and its proteolytic cleavage in HIV-1 structural maturation and infectivity. *J. Virol.* **86**, 13708–13716 (2012).
42. Kingston, R. L. *et al.* Structure and self-association of the Rous sarcoma virus capsid protein. *Structure* **8**, 617–628 (2000).
43. Campos-Olivas, R., Newman, J. L. & Summers, M. F. Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. *J. Mol. Biol.* **296**, 633–649 (2000).
44. Mortuza, G. B. *et al.* High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature* **431**, 481–485 (2004).
45. Cornilescu, C. C., Bouamr, F., Yao, X., Carter, C. & Tjandra, N. Structural analysis of the N-terminal domain of the human T-cell leukemia virus capsid protein. *J. Mol. Biol.* **306**, 783–797 (2001).
46. Khorasanizadeh, S., Campos-Olivas, R., Clark, C. & Summers, M. Sequence-specific <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift assignment and secondary structure of the HTLV-I capsid protein. *J. Biomol. NMR* **14**, 199–200 (1999).





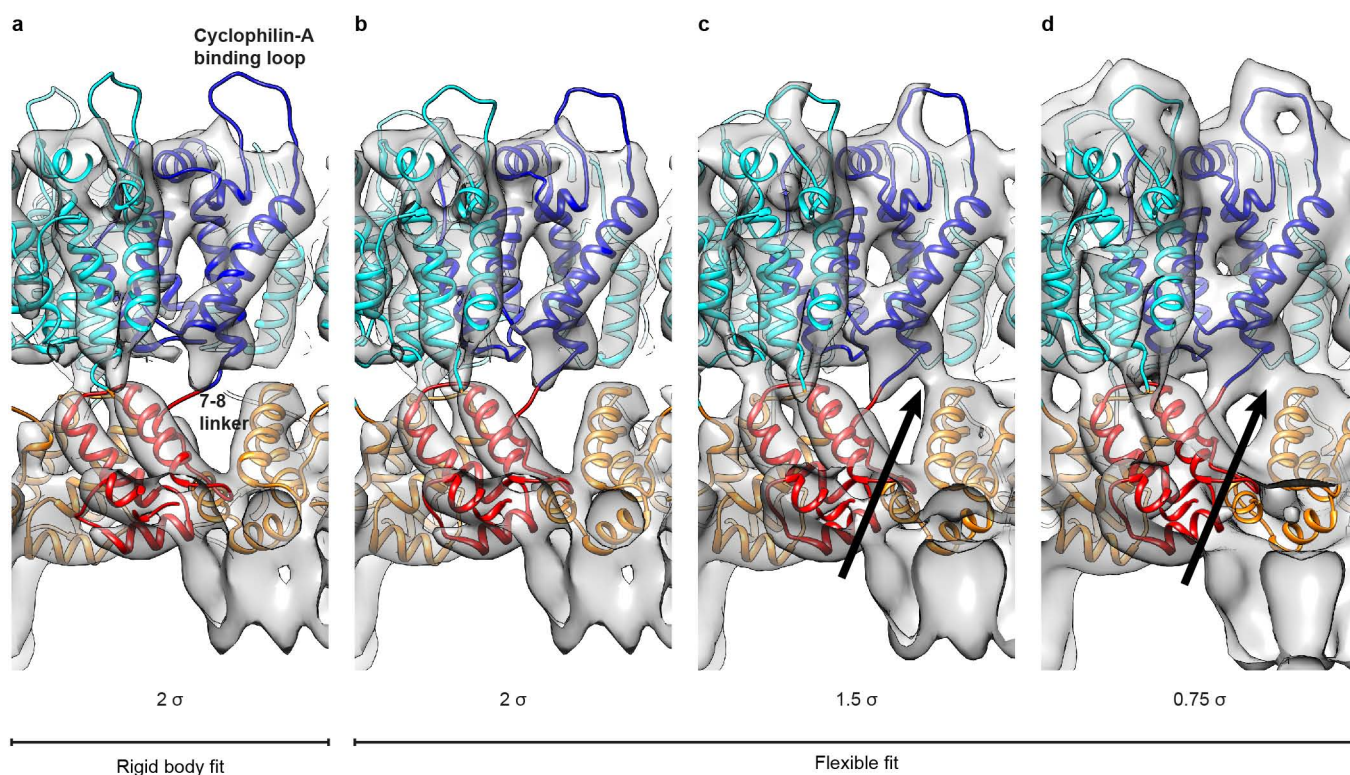
**Extended Data Figure 1 | Characterization of virus preparations used in this study.** Samples of Optiprep gradient (OP)-purified particles were separated by SDS–polyacrylamide gel electrophoresis. Proteins were visualized by silver staining or immunoblot, respectively, as indicated. For immunoblot analysis, proteins were transferred to nitrocellulose membranes by semi-dry blotting. Membranes were probed with polyclonal antiserum raised against recombinant HIV-1 CA (**a**) or M-PMV NC (**c**), respectively. Bound antibodies were visualized by quantitative immunodetection on a LiCor Odyssey imager, using secondary antibodies and protocols according to the manufacturer's instructions. **a**, HIV-1 particles prepared in the presence of 5  $\mu$ M APV. Note

that residual processing of Gag has occurred due to incomplete protease inhibition. The main additional Gag product ( $\sim$ 50 kDa) corresponds to Gag lacking the C-terminal p6-region; we have previously shown that cleavages downstream of SP1 do not disrupt the immature Gag lattice<sup>41</sup>. **b**, Immature HIV-1 particles prepared from cells transfected with pNL4-3 (PR-, D25A) compared with a standard of purified Gag. Note that Gag is completely uncleaved in this case. **c**, Immature protease defective M-PMV particles purified from cells transfected with plasmid pSHRM15 (D26N). Positions of molecular mass standards (in kilodaltons) are indicated.



**Extended Data Figure 2 | Cryo-electron tomography and subtomogram averaging reconstruction of the immature HIV-1 lattice.** **a**, Within pleiotropic HIV particles the polyprotein Gag forms incomplete hexameric lattices of heterogeneous shapes. Tomographic slices of two representative viruses are shown. Underneath the membrane, ordered protein density corresponding to capsid can be seen. In the middle and lower panels, virus particles are displayed with their respective output lattice maps as derived from subtomogram averaging. The colour of the hexagons denotes the cross-correlation coefficient (CCC) of the alignment ranging from red (low CCC (0.04)) to green (high CCC (0.19)). In the lower panel, the view has been tilted to reveal regions with no lattice. Areas with no protein density underneath the membrane are devoid of a hexameric lattice. Scale bar, 50 nm. **b**, FSC between two half-data sets (each of which was independently aligned and averaged starting from independent references) of immature HIV treated with APV

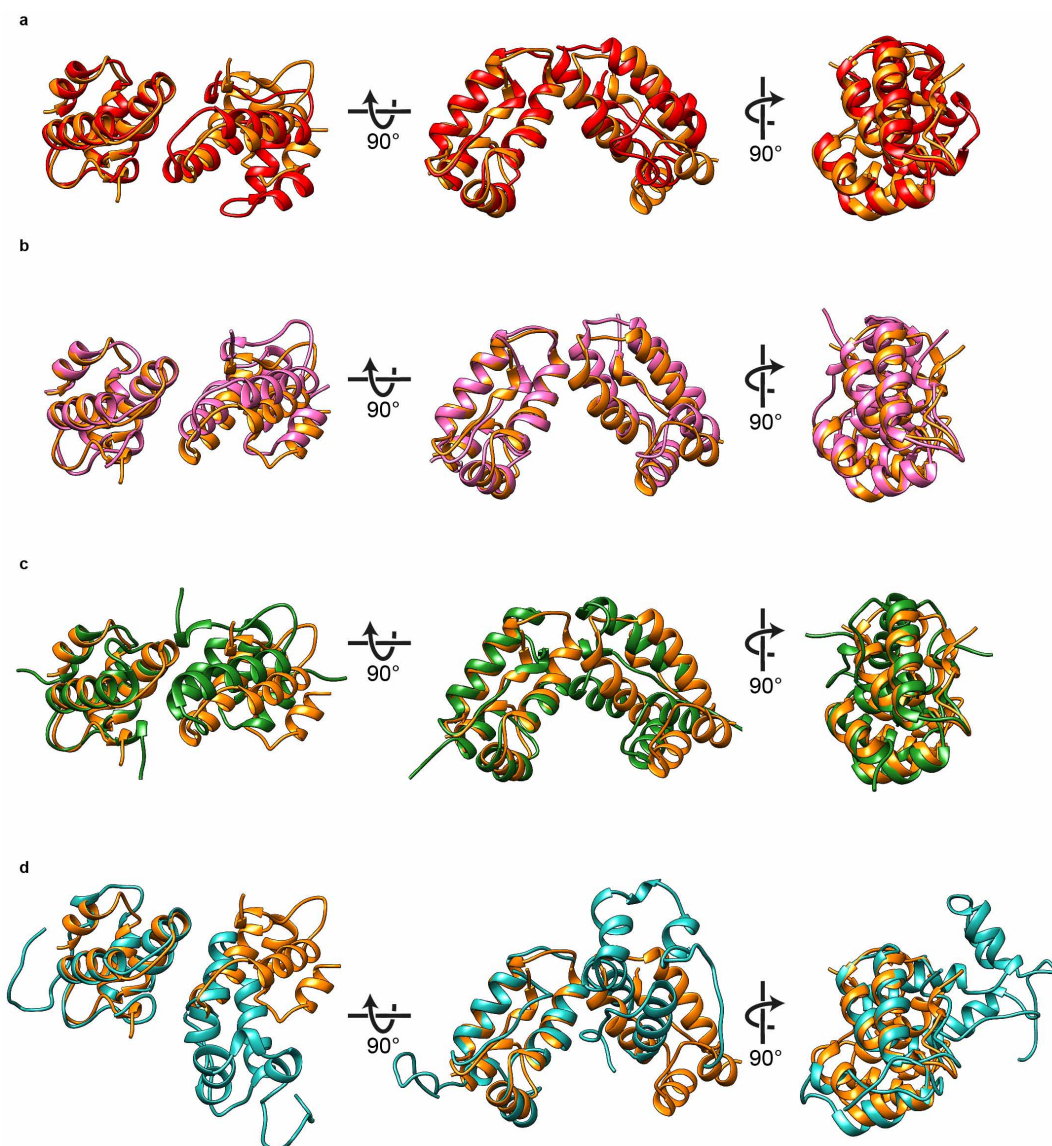
(blue line) showing a resolution of 9.8/8.8 Å at the 0.5/0.143 criterion, respectively. The equivalent FSC for a lower-resolution structure of the protease defective HIV (D25A) sample (red line) shows a resolution of 12.6/10.9 Å at the 0.5/0.143 criterion. **c**, Central orthoslice through the final average of the protease-inhibitor-treated HIV reconstruction (no inverse B-factor was applied). Ordered density is only observed in the CA-SP1 region. The dashed rectangle indicates the region shown in **d**. **d**, Orthoslice through the CA-SP1 region in the final average of HIV-1 + APV (left, corrected with an inverse B-factor of  $-1200 \text{ Å}^2$ ), the same structure filtered to 13 Å (middle) and of the protease defective HIV-1 (D25A) preparation filtered to 13 Å (right), indicating that protease-inhibited and protease mutant structures have the same domain arrangement. The dashed lines indicate the positions of horizontal orthoslices in **e** and **f**. Scale bar, 50 Å.



**Extended Data Figure 3 | Molecular dynamics flexible fitting of high-resolution HIV-1 CA structures into the electron microscopy density.** **a**, The rigid body fit used as a starting model for flexible fitting superimposed onto the electron microscopy map. **b–d**, The final flexible fit superimposed onto the electron microscopy map viewed at three different isosurface thresholds. The flexible fitting resulted in only minor movements relative to the initial rigid-body fit, implying that individual CA domains do not undergo large

changes in structure between the isolated and assembled protein domains. At lower isosurface thresholds, densities corresponding to the cyclophilin-A binding loop and the 7–8 linker can be seen (positions annotated in **a**). One of the positions at which CA-NTD and CA-CTD contact one another is marked by an arrow. Isosurface values are indicated in the figure (threshold value is  $\sigma$  away from the mean).

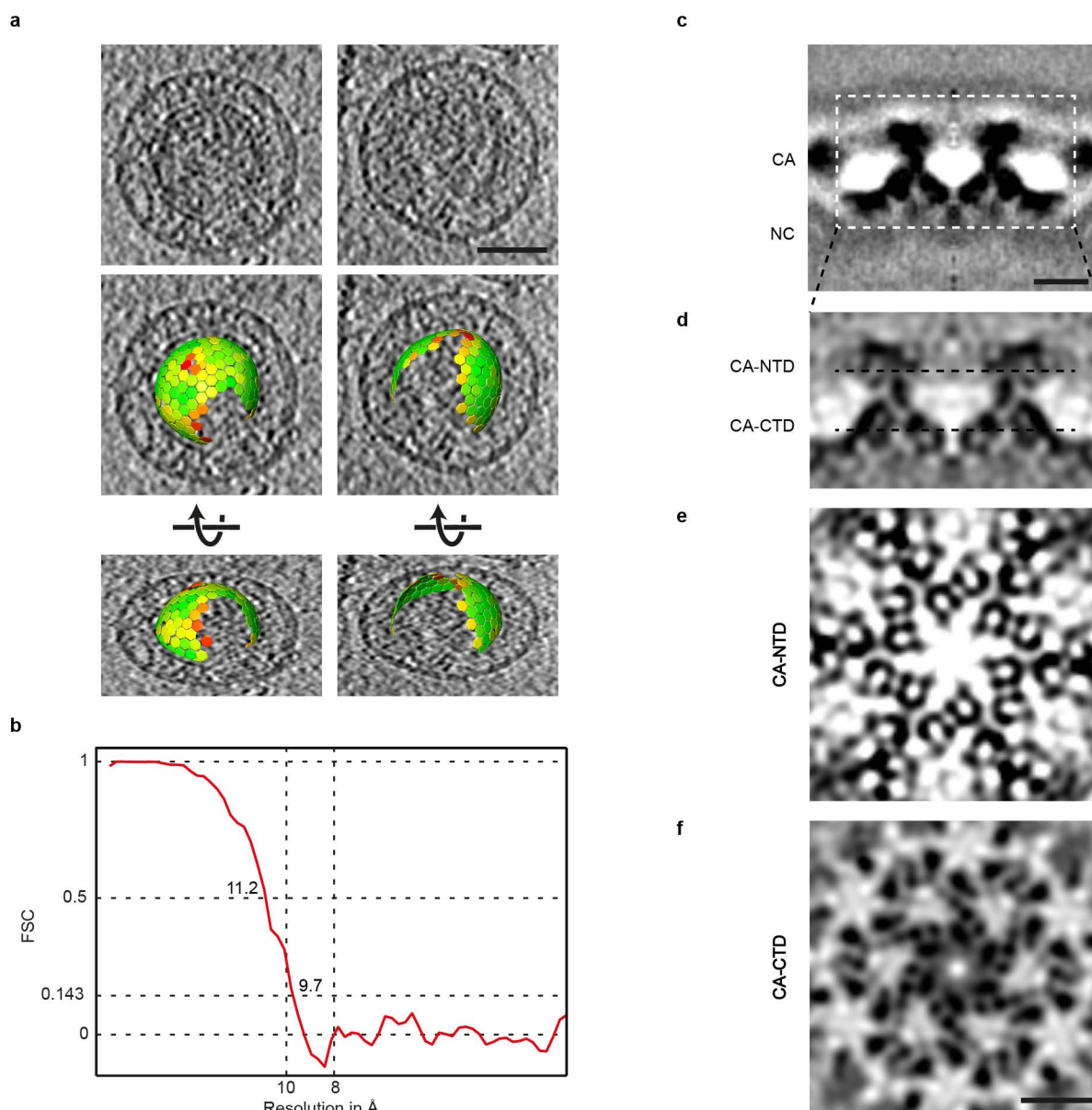




**Extended Data Figure 4 | Comparison of different retroviral CA-CTD dimer structures.** **a–d**, Comparison of different CA-CTD dimers aligned against the immature flexibly fitted HIV-1 CA-CTD dimer obtained in this study (orange). **a**, PDB 3DS2 (red, the crystal structure that most closely resembles the immature CA-CTD); **b**, the structure recently obtained by fitting a crystal dimer into *in vitro*-assembled HIV-1 tubes<sup>7</sup> (pink); **c**, the immature

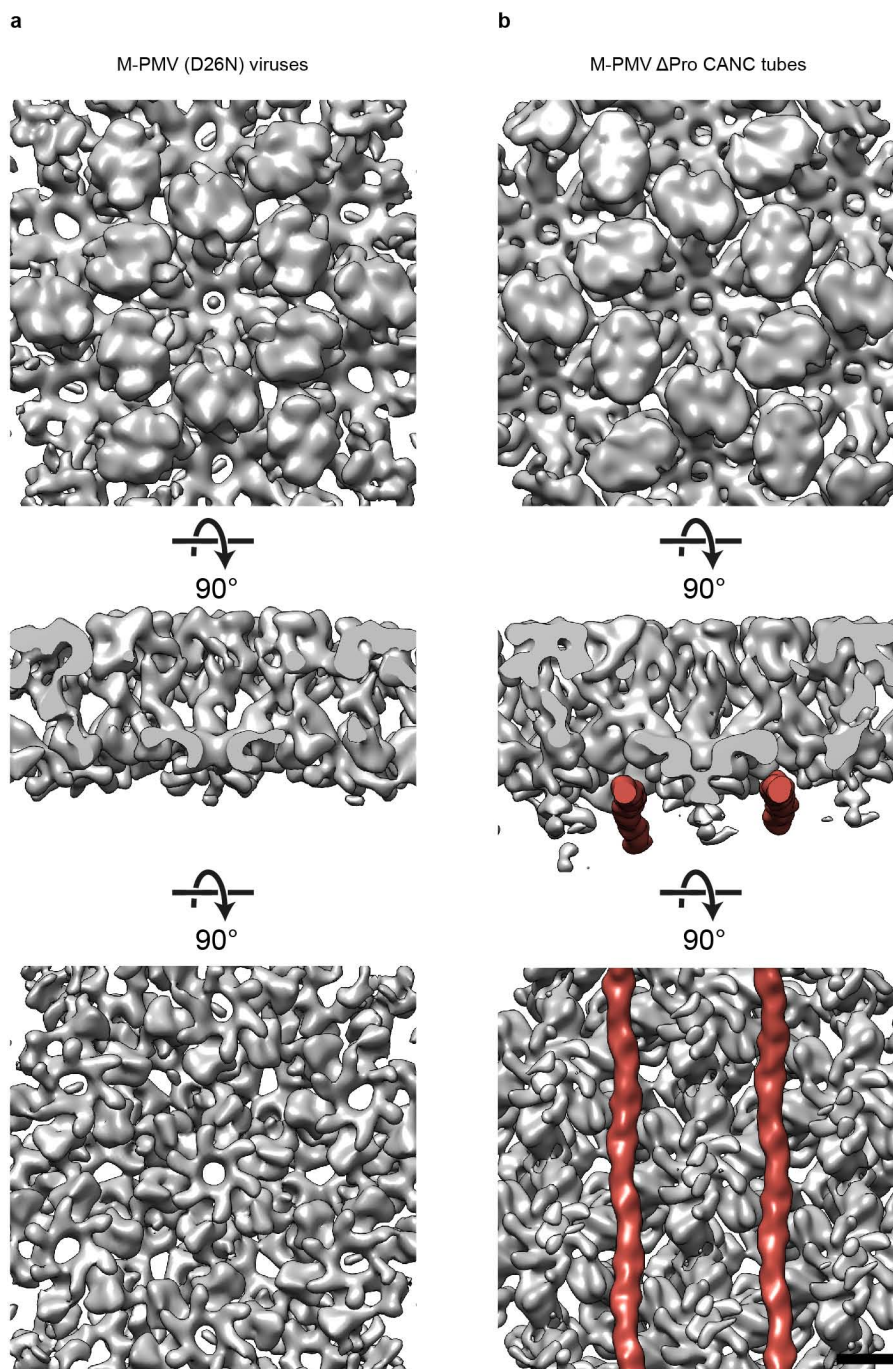
M-PMV dimer based on a homology model fitted into the immature M-PMV electron microscopy density generated in this study (green); **d**, a mature HIV-1 CA-CTD dimer from PDB 3J34 (ref. 6) (light blue, chain A and f). The backbone root mean squared deviations between the superimposed structures are 6.8 Å (**a**), 3.2 Å (**b**), 6.4 Å (**c**) and 17.9 Å (**d**).





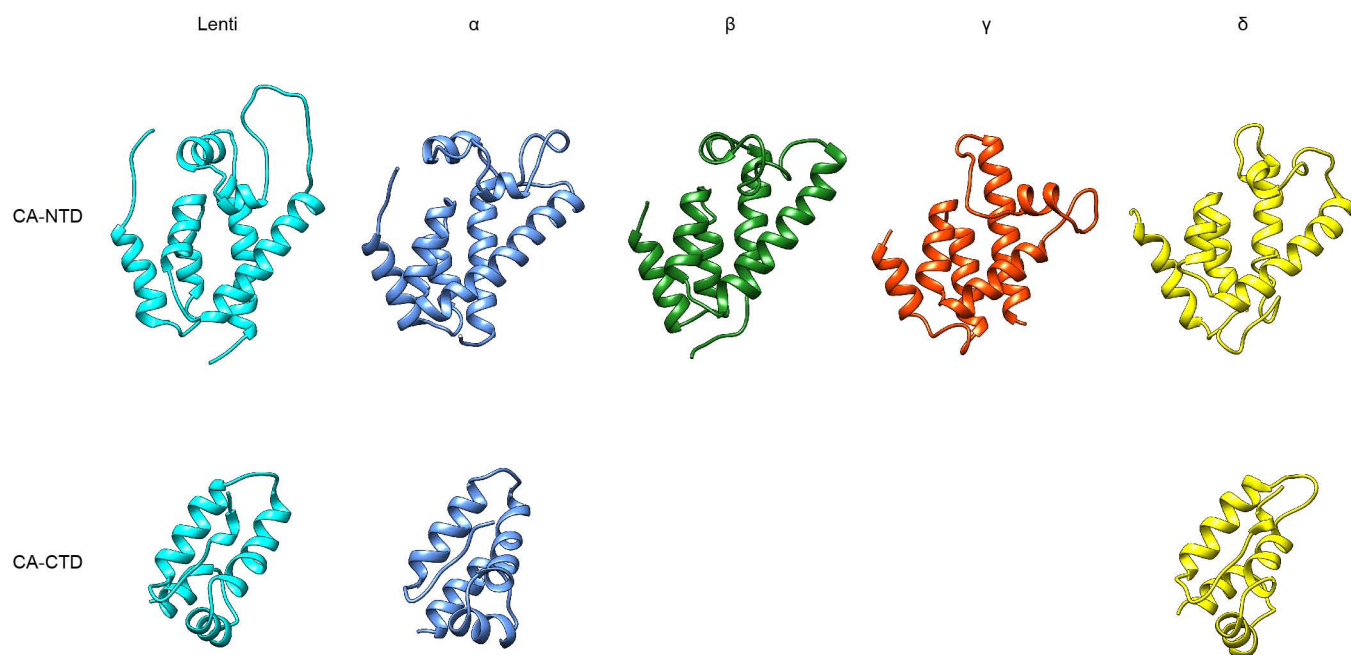
**Extended Data Figure 5 | Cryo-electron tomography and subtomogram averaging reconstruction of immature M-PMV (D26N) viruses.** **a**, Slices through two tomograms containing immature M-PMV (D26N) particles. Using subtomogram averaging, the position and the arrangement of the hexameric capsid lattice can be resolved (middle and lower panels). See Extended Data Fig. 2 for explanations. CCC values range from 0.05 to 0.14. Scale bar, 50 nm. **b**, FSC of the two independently aligned and averaged half-data sets of the M-PMV (D26N) sample. The resolution corresponds to

11.2 and 9.7 Å at the 0.5/0.143 criterion, respectively. **c**, Side view orthoslice through the final average of the protease defective M-PMV reconstruction (no inverse B-factor was applied). No ordered densities are observed except for the CA region. The dashed rectangle indicates the region shown in **d**. **d**, Orthoslice through the CA region in the final average corrected with a B-factor of  $-1,000 \text{ Å}^2$ . The dashed lines indicate the positions of horizontal orthoslices in **e** and **f**. Scale bar, 50 Å.



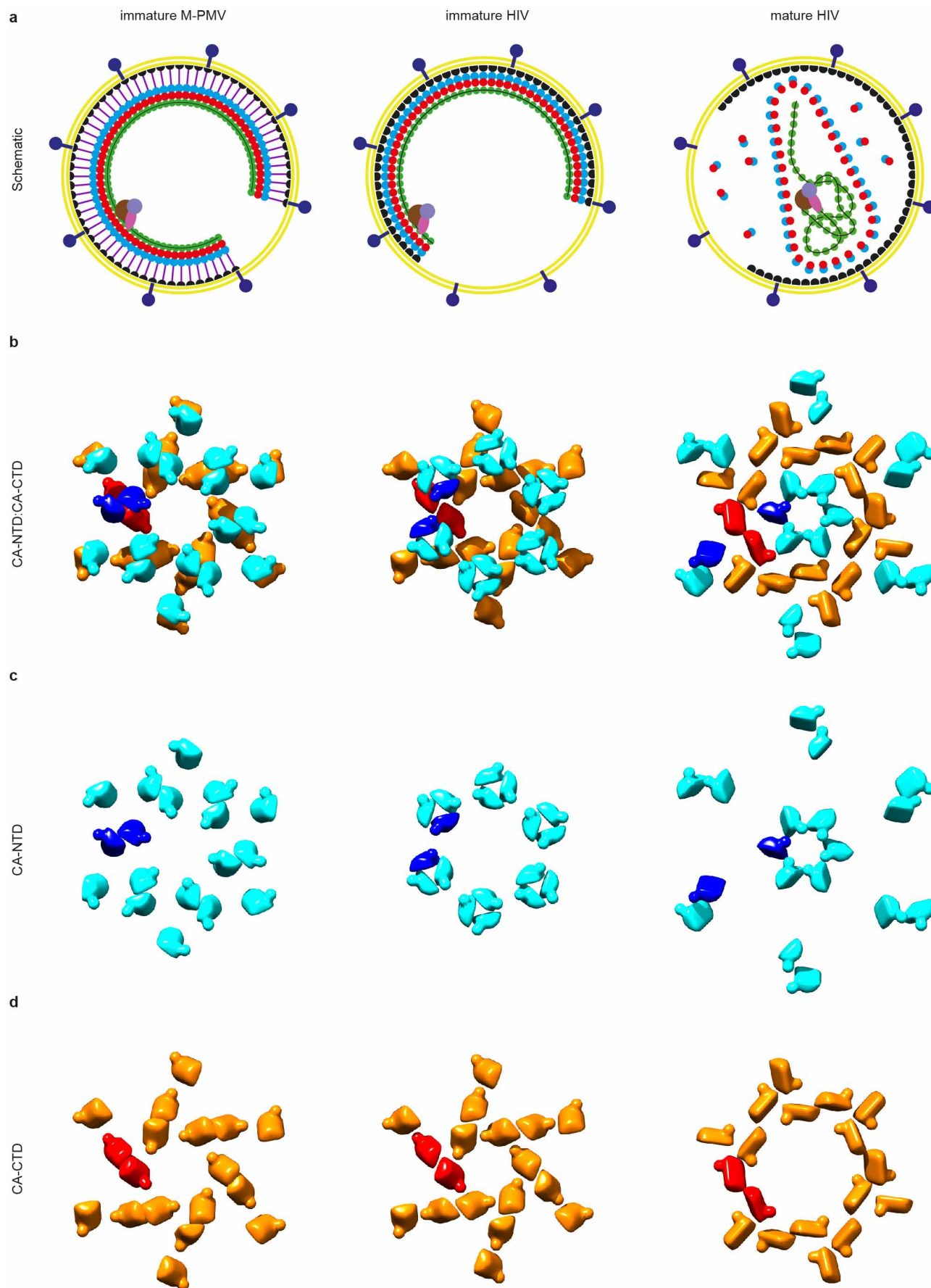
**Extended Data Figure 6 | Comparison of immature protease defective M-PMV (D26N) particles and M-PMV  $\Delta$ Pro CANC tubes.** **a**, An isosurface representation of the structure generated from M-PMV (D26N) virus particles, as presented in this study. The structure is shown from the outside of the virus (top), in a horizontal slice (middle) and from the inside of the virus (bottom). **b**, For comparison, equivalent views of the structure from

*in vitro*-assembled M-PMV  $\Delta$ Pro CANC tubular arrays (EMD2089)<sup>2</sup> are shown. Both structures have been filtered to 10 Å and the threshold was set to  $2\sigma$  away from the mean. The additional density observed for the M-PMV  $\Delta$ Pro CANC tube structure (coloured red in the middle and bottom panel) is thought to be a nucleic acid structure that is ordered in the tubular arrays<sup>2</sup>. Scale bar, 25 Å.



**Extended Data Figure 7 | Structural conservation of retroviral CA domains.** Known atomic structures of CA-NTD and CA-CTD domains from the lentivirus<sup>13,14</sup>,  $\alpha$ <sup>42,43</sup>,  $\beta$ <sup>40</sup>,  $\gamma$ <sup>44</sup> and  $\delta$ <sup>45,46</sup> retrovirus families. The tertiary structure of CA domains is highly conserved between different retroviruses.







**Extended Data Figure 8 | Comparison of CA domain arrangements in immature M-PMV, immature HIV and mature HIV.** **a**, Two-dimensional schematic representations of immature M-PMV, immature HIV and mature HIV particles. Retroviruses bud from infected cells in an immature form. The Gag polyprotein is radially arranged underneath the host-derived plasma membrane (yellow) and includes the membrane associated matrix domain (MA, black), the bipartite capsid domain (CA, blue and red) and the nucleic-acid-bound nucleocapsid (NC, green). M-PMV possesses two additional domains positioned between MA and CA, termed pp24 and p12 (purple), leading to a bigger spatial separation of CA and MA compared with HIV. The viral protease cleaves the Gag polyprotein at defined positions, triggering maturation. This process leads to a rearrangement of the domains, giving rise to

the mature, infectious virus. **b–d**, Schematic diagram representing the arrangement of the CA-NTD and CA-CTD domains within immature M-PMV and HIV and within mature HIV. CA-NTD and CA-CTD molecules are represented by cyan/blue- and orange/red-coloured solids, respectively. The solids are positioned at the exact positions and orientations at which the high-resolution structures fitted into the electron microscopy densities. For CA-NTD the N terminus and for CA-CTD the C terminus are represented as spherical extensions. The shown schematics were generated in UCSF Chimera by defining the translational and rotational matrix of each fitted CA domain, and applying it to the solid representations. **b**, Both CA-NTD (cyan/blue) and CA-CTD (orange/red) domains are shown; **c**, CA-NTD only; **d**, CA-CTD only.

Extended Data Table 1 | Data processing statistics

	Immature protease inhibitor treated (APV) HIV viruses	Immature protease defective HIV (D25A) viruses	Immature protease defective M-PMV (D26N) viruses
<b>Microscope</b>	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
<b>Voltage (keV)</b>	200	300	200
<b>Detector</b>	Gatan multiscan 795 CCD camera	FEI Falcon II Direct detector	Gatan multiscan 795 CCD camera
<b>Energy-filter</b>	GATAN GIF2002	None	GATAN GIF2002
<b>Å/pixel</b>	2.025	2.281	2.025
<b>Defocus range (microns)</b>	-1.2 to -4.0	-2.0 to -5.5	-1.5 to -4.5
<b>Total Dose (electrons/Å<sup>2</sup>)</b>	40	60	40
<b>Tomograms reconstructed</b>	70	33	39
<b>Viruses extracted</b>	224	116	84
<b>Asymmetric units Set A</b>	94,986	33,762	34,818
<b>Asymmetric units Set B</b>	99,744	29,352	42,702
<b>Final resolution (0.5/0.143 FSC) in Å</b>	9.8/8.8	12.6/10.9	11.2/9.7

Summary of image acquisition and processing statistics for the data sets discussed in this study. Note that each reconstruction was performed on two independent half-data sets.

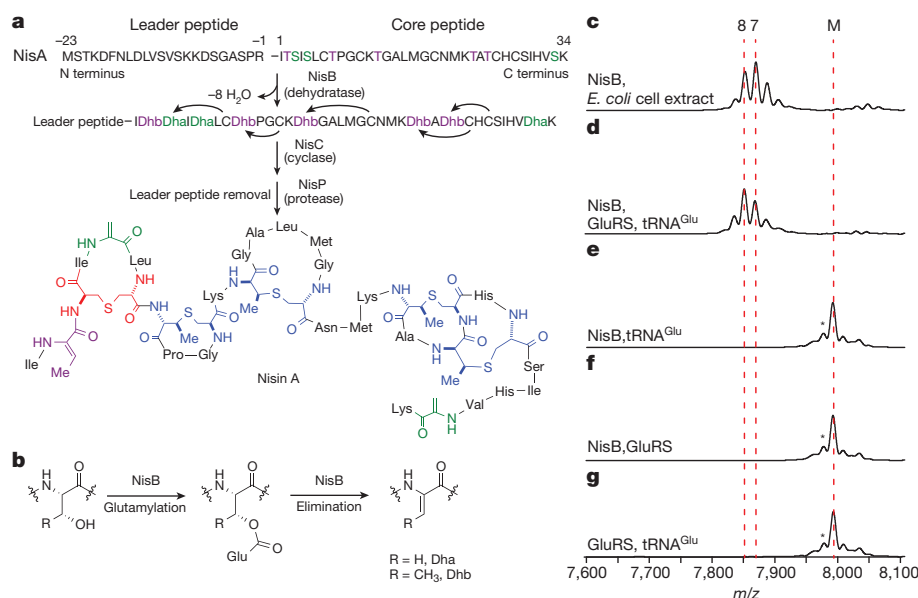
# Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB

Manuel A. Ortega<sup>1\*</sup>, Yue Hao<sup>1\*</sup>, Qi Zhang<sup>2</sup>, Mark C. Walker<sup>2</sup>, Wilfred A. van der Donk<sup>1,2</sup> & Satish K. Nair<sup>1,3</sup>

Lantibiotics are a class of peptide antibiotics that contain one or more thioether bonds. The lantibiotic nisin is an antimicrobial peptide that is widely used as a food preservative to combat food-borne pathogens<sup>1</sup>. Nisin contains dehydroalanine and dehydrobutyrine residues that are formed by the dehydration of Ser/Thr by the lantibiotic dehydratase NisB (ref. 2). Recent biochemical studies revealed that NisB glutamylates Ser/Thr side chains as part of the dehydration process<sup>3</sup>. However, the molecular mechanism by which NisB uses glutamate to catalyse dehydration remains unresolved. Here we show that this process involves glutamyl-tRNA<sup>Glu</sup> to activate Ser/Thr residues. In addition, the 2.9-Å crystal structure of NisB in complex with its substrate peptide NisA reveals the presence of two separate domains that catalyse the Ser/Thr glutamylation and glutamate elimination steps. The co-crystal structure also provides insights into substrate recognition by lantibiotic dehydratases. Our findings demonstrate an unexpected role for aminoacyl-tRNA in the formation of dehydroamino acids in lantibiotics, and serve as a basis for the functional characterization of the many lantibiotic-like dehydratases involved in the biosynthesis of other classes of natural products.

Bacterial resistance to currently used antibiotics is a growing health threat. A potential solution to this emerging problem is the development of new antibiotics with multiple modes of action that would challenge bacterial resistance mechanisms. For instance, the lantibiotic nisin has been used worldwide in the food industry for more than 40 years without substantial development of resistance<sup>1,4</sup>. This unique property is thought to be a consequence of nisin's dual mode of action: pore formation in bacterial cell membranes and stalling of peptidoglycan biosynthesis by sequestering the cell-wall precursor lipid II<sup>5–7</sup>.

Lantibiotics are lanthionine-containing antimicrobial peptides<sup>8</sup>. Nisin is generated from a ribosomally synthesized linear precursor peptide termed NisA<sup>9</sup>. The dehydratase NisB then dehydrates eight serines and threonines in the NisA core region, yielding dehydroalanine (Dha) and dehydrobutyrine (Dhb) residues, respectively (Fig. 1a)<sup>2</sup>. The cyclase NisC subsequently catalyses the formation of five lanthionine and methylanthionine crosslinks by the nucleophilic addition of cysteinyl thiols to Dha and Dhb, respectively (Fig. 1a)<sup>10</sup>. In the final maturation step, the lantibiotic protease NisP removes a leader peptide, which is important for recognition by NisB and NisC, to yield the mature lantibiotic<sup>11</sup>.



**Figure 1 | Biosynthesis of the lantibiotic nisin.** **a**, Post-translational modifications involved in nisin biosynthesis. **b**, NisB dehydrates Ser/Thr residues in NisA by glutamylation, forming Dha and Dhb, respectively. NisC catalyses the formation of lanthionines (red) and methylanthionines (blue). NisP removes the leader peptide. Negative numbers represent the position of an amino acid in the leader peptide with respect to the core region. **c–g**, MALDI-TOF MS analysis of His<sub>6</sub>-NisA incubated with Glu, ATP and His<sub>6</sub>-NisB in the

presence of *E. coli* cell extract (**c**), His<sub>6</sub>-GluRS and tRNA<sup>Glu</sup> (**d**), tRNA<sup>Glu</sup> (**e**), His<sub>6</sub>-GluRS (**f**) or GluRS and tRNA<sup>Glu</sup> in the absence of His<sub>6</sub>-NisB (**g**). M, unmodified His<sub>6</sub>-NisA (*m/z* 7,993, calc. *m/z* 7,996); 7, sevenfold dehydrated His<sub>6</sub>-NisA (*m/z* 7,870, calc. *m/z* 7,870); 8, eightfold dehydrated His<sub>6</sub>-NisA (*m/z* 7,853, calc. *m/z* 7,852). The asterisk indicates a peak resulting from laser-induced deamination (Extended Data Fig. 6).

<sup>1</sup>Department of Biochemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. <sup>2</sup>Department of Chemistry and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. <sup>3</sup>Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA.

\*These authors contributed equally to this work.

During the 26 years since the characterization of the first lantibiotic gene cluster<sup>12</sup>, the mechanism by which lantibiotic dehydratases (LanB) introduce dehydroamino acids into class I lantibiotics such as nisin has remained enigmatic. Recently, NisB was shown to dehydrate NisA via an unprecedented glutamylation mechanism (Fig. 1b)<sup>3</sup>. However, NisB was only active in the presence of an unknown component within *Escherichia coli* cell extract<sup>3</sup>. Here we identify glutamyl-tRNA<sup>Glu</sup> as the key component needed to catalyse the formation of dehydroamino acids in class I lantibiotics. In addition we report the co-crystal structure of NisA bound to NisB, which provides key information on the glutamyl-tRNA<sup>Glu</sup>-dependent esterification of Ser/Thr residues in NisA and offers the first insights into leader peptide binding and substrate recognition by lantibiotic dehydratases.

In the previously proposed dehydration mechanism<sup>3</sup>, glutamate needs to be activated before the formation of an ester linkage with the side chain of Ser/Thr residues in NisA. To identify the required component for activation, anion-exchange chromatographic fractions of *E. coli* cell extracts were analysed for supporting NisB-catalysed dehydration of NisA by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Extended Data Fig. 1). The  $A_{260}/A_{280}$  ratio for the fraction supporting NisB activity was 1.7, suggesting the presence of nucleic acids. Treatment of *E. coli* cell extract with DNase did not prevent the NisB-catalysed dehydration of NisA, but treatment with RNase abolished activity (Extended Data Fig. 1).

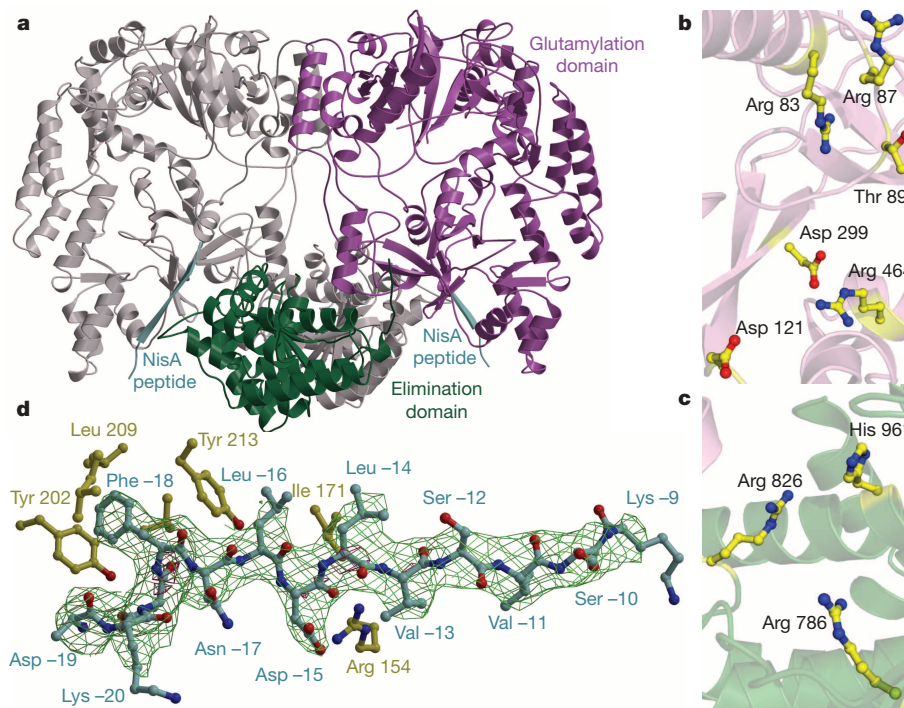
The requirements for glutamate and RNA suggested the possibility of a glutamyl-tRNA<sup>Glu</sup>-dependent dehydration process. We therefore cloned, expressed and purified glutamyl-tRNA synthetase (GluRS) and tRNA<sup>Glu</sup> from *E. coli*. MALDI-TOF MS analysis of NisA after incubation with NisB, Glu, ATP, GluRS and tRNA<sup>Glu</sup> revealed up to eight dehydrations in NisA (Fig. 1c, d). No activity was observed when GluRS, tRNA<sup>Glu</sup> or NisB was omitted from the reaction assay (Fig. 1e–g). The inability of NisB to dehydrate NisA in the absence of GluRS (Fig. 1e) demonstrates that NisB does not aminoacylate tRNA<sup>Glu</sup> but instead uses the glutamyl-tRNA<sup>Glu</sup> synthesized *in situ* by GluRS. Dehydration assays

with purified glutamyl-tRNA<sup>Glu</sup> and subsequent MALDI-TOF MS analysis confirmed this conclusion and also showed that ATP was not required for activity (Extended Data Fig. 2). This observation suggests that dehydration of Ser/Thr residues located at different positions within NisA is not driven by the consumption of energy.

The observed requirement for *E. coli* glutamyl-tRNA<sup>Glu</sup> raised the question whether the nisin-producing organism, *Lactococcus lactis*, employs the same tRNA-dependent biosynthetic strategy. MALDI-TOF MS analysis of NisA treated with NisB, Glu, ATP and both GluRS and total RNA from *L. lactis* revealed up to eight dehydrations in NisA (Extended Data Fig. 3), confirming that NisB can use glutamyl-tRNA<sup>Glu</sup> from *L. lactis*.

To obtain structural insights into the dehydration process, we determined the co-crystal structure of NisB in complex with its substrate peptide NisA to 2.9 Å resolution (Extended Data Table 1). The overall structure of the 117-kDa dehydratase reveals a dimer with a bifurcated cleft-like structure in the centre (Extended Data Fig. 4). Residues spanning Ser 5–Ile 706 comprise a multi-domain amino-terminal region, and Arg 716–Glu 990 form a single carboxy-terminal domain (Fig. 2a). Residues covering Ile 706–Arg 716 and Val 893–Gly 919 were not modelled in the final structure, because of a lack of electron density.

NisB mutations previously shown to interfere with either glutamylation or elimination activity map to distinct clusters in the N-terminal and C-terminal regions in the structure, respectively (Extended Data Fig. 4). Alanine substitutions of NisB residues that abolished the glutamylation of NisA but retained *in vitro* glutamate elimination activity<sup>3</sup> are all within a 10 Å radius in the N-terminal region (Fig. 2b). This 700-residue N-terminal region of NisB can thus be demarcated as the glutamylation domain. Similarly, alanine-scanning studies in NisB demonstrated that Arg 786, Arg 826 and His 961 are important for glutamate elimination but not for glutamylation activity<sup>3</sup>. These residues, as well as others important for dehydration on the basis of mutagenesis data<sup>3</sup> (Extended Data Fig. 5), cluster to a small site in the C-terminal domain (Fig. 2c), and consequently, the C-terminal 300-residue domain of NisB



**Figure 2 | Crystal structure of the lantibiotic dehydratase NisB.** **a**, Overall structure of the NisB homodimer in complex with its substrate peptide NisA, showing the disposition of the glutamylation (purple) and glutamate elimination (green) domains; the other monomer is shown in grey. The NisA peptide is shown in cyan. **b**, **c**, Residues in NisB essential for either the

glutamylation (**b**) or glutamate elimination activities (**c**) are clustered in the crystal structure. **d**, Simulated annealing omit difference Fourier map ( $F_o - F_c$ ) contoured to  $2.5\sigma$  of residues Lys -20 to Lys -9 within the leader peptide of NisA. NisB residues involved in binding of the NisA leader peptide (cyan) are shown in yellow.



catalyses the glutamate elimination step. The boundaries for the two domains correspond to two proteins of roughly 800 residues and 300 residues encoded in the biosynthetic clusters of thiopeptides<sup>13</sup>, suggesting that they conduct aminoacylation and elimination, respectively, to achieve the dehydration of their substrate peptides.

RNA binding proteins often have basic patches that mediate the binding of their cognate RNAs by means of electrostatic interactions<sup>14</sup>. Mapping the electrostatic potential onto the surface of NisB reveals the presence of a densely basic cavity that maps to the inner face of the cleft-like structure (Extended Data Fig. 4). The dimensions of the basic cavity are suitable for accommodating the characteristic L shape of polyanionic tRNA. On the basis of limited structural homology of this NisB region to the *X. laevis* double-stranded RNA binding protein A complexed to its cognate RNA (PDB 1DI2) and using a structure of bacterial tRNA from *Thermus thermophilus* (PDB 1N78), a docking pose was derived for NisB binding to tRNA<sup>Glu</sup> (Extended Data Fig. 4). This model results in placement of the 3' end of the tRNA in close proximity to the pocket defined by the NisB residues that are critical for glutamylation activity, providing additional support for the model.

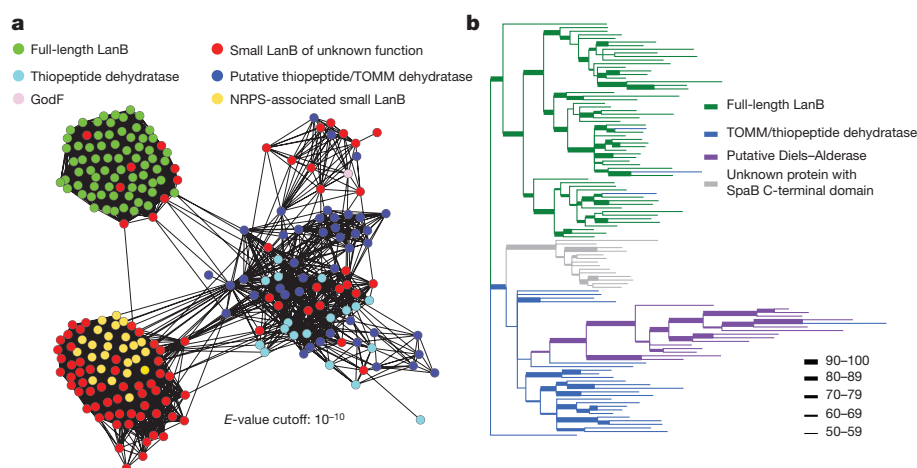
Structure-based similarity searches against the Protein Data Bank failed to identify obvious homologous folds within the NisB structure. However, searches using smaller domains of NisB identified several structural homologues (Extended Data Fig. 4). A small domain spanning residues Phe 136–Ile 216 resembles the N-terminal domain of TruD, a heterocyclase involved in the biosynthesis of cyanobactins<sup>15</sup>, another class of ribosomally synthesized and post-translationally modified peptides (RiPPs) that is made in a leader-peptide-dependent manner<sup>16</sup>. A second region within the C terminus of NisB (Phe 734–Tyr 820) is similar to the LsrG protein that performs the epimerization of activated quorum-sensing molecules<sup>17</sup>. However, despite the structural homology and similar chemistry (abstraction of a relatively acidic proton adjacent to a carbonyl group), residues important for LsrG catalysis<sup>17</sup> are not conserved in NisB. Instead, residues important for glutamate elimination in NisB are located between the LsrG-like region and a second subdomain of the glutamate elimination site (Extended Data Fig. 5). Indeed, the crystal structure of NisB allowed us to identify an additional conserved positively charged residue in this pocket, Arg 784. As expected, the NisB-R784A mutant resulted in the accumulation of glutamylated NisA intermediates (Extended Data Fig. 5), providing additional support for our assignment of this region as the active site for glutamate elimination.

Clear and continuous electron density was observed for the NisA leader sequence spanning residues Lys –9 to Lys –20 (Fig. 2d). Beyond Lys –9, moving towards the C-terminal core region, the electron density diminishes, probably as a result of NisA conformational flexibility. The NisA leader peptide binds to NisB residues Leu 166–Asn 175 as an antiparallel  $\beta$ -strand. The site of binding combined with the extended conformation of the leader peptide ensures that the core region reaches

the active site for glutamylation. NisA, like most class I lantibiotic precursor peptides, contains a F-D/N-L-N/D conserved motif within the leader peptide that is important for substrate recognition by NisB<sup>18–20</sup>. The structure reveals that Phe –18 of NisA in this motif binds to NisB within a hydrophobic cage composed of Val 176, Val 198, Tyr 202, Leu 209 and Tyr 213. Similarly, Leu –16 of NisA is clustered inside a NisB pocket consisting of residues Ile 171, Tyr 213 and Leu 217. Last, Asp –15 of NisA is within hydrogen-bonding distance of Arg 154 (Fig. 2d). The interactions observed in the crystal structure provide an explanation of why mutations within the NisA FNLD sequence greatly decrease NisB binding<sup>19,20</sup>. NisA binds to the region in NisB that is homologous to the N-terminal domain of TruD (Extended Data Fig. 4), indicating that some of the different classes of biosynthetic enzymes involved in the processing of RiPPs probably use similar structural elements for substrate binding.

The presence of the leader-peptide-binding site on the N-terminal domain raised the question of whether the glutamate elimination activity requires the leader peptide. To gain further insight, we tested NisB-catalysed glutamate elimination activity on purified glutamylated NisA core peptide in the absence of the leader peptide. MALDI–TOF MS analysis after incubation with NisB revealed the presence of dehydrated NisA core peptide (Extended Data Fig. 6). The leader peptide therefore does not seem to be crucial for glutamate elimination, suggesting that the local structure of glutamylated Ser/Thr is sufficient for substrate recognition by the elimination domain.

The data presented here provide explanations for several long-standing questions. NisB substrate recognition is mediated by hydrophobic interactions between NisB and the FNLD motif in NisA. The presence of this motif in many class I lantibiotic precursor peptides suggests that a similar recognition mechanism is used by their respective lantibiotic dehydratases. In addition, the NisB-catalysed dehydration of NisA has been shown to proceed with general N-to-C-terminal directionality<sup>21,22</sup>. This observation, combined with the demonstration that NisB dehydrates NisA in a distributive manner (Extended Data Figs 1–3) and with the leader-peptide-binding site identified in NisB, suggests that directionality is achieved by the distance between the leader-peptide-binding site and the active site for glutamylation. Residues closer to the leader peptide (N terminus) have a higher likelihood of accessing the glutamylation active site than residues located at the C terminus. The presence of a distinct leader-peptide-binding site also explains why Ser/Thr residues within the leader peptide are never dehydrated and why there is a minimum distance requirement from the F-D/N-L-N/D motif to the first dehydrated amino acid in class I lantibiotics<sup>19,23</sup> (that is, the distance required to reach the glutamylation site). After glutamylation, the flexible core region reaches the active site for glutamate elimination, where the glutamate adduct is then eliminated. This flexibility can account for the diminished electron density observed for the NisA C-terminal portion.



**Figure 3 | Protein similarity map and phylogenetic tree analysis of various lantibiotic dehydratases and homologues.** **a**, Protein similarity map of the glutamylation domain (Lant\_dehyd\_C, PF04738) present in various lantibiotic dehydratases and homologues. Each node in the network represents a protein sequence, and each edge represents sequences with BLASTP *E* values below the indicated cutoff. NRPS, non-ribosomal peptide synthesis; TOMM, thiazole/oxazole-modified microcin. **b**, Phylogenetic tree analysis of the glutamate elimination domain (SpaB\_C, PF14028) present in various lantibiotic dehydratases and homologues. Several clades can be identified on the basis of the proposed function and natural product class. Only splits with a posterior probability higher than 50% are shown.

LanBs are widely distributed among Gram-positive bacteria. A protein similarity map of the glutamylation domain of LanB proteins (Supplementary Table 1) in the databases identified several subfamilies (Fig. 3a). One large group is formed predominantly by sequences that contain the glutamylation and elimination domains in one polypeptide (full-length LanB; green in Fig. 3a). Another less-well-defined group is formed by dehydratases involved in the biosynthesis of thiopeptides (cyan in Fig. 3a) and goadsporin (pink in Fig. 3a). In these LanBs the glutamylation and elimination domains are in separate polypeptides. Whether these enzymes do indeed use glutamate or perhaps another amino acid for the activation of Ser/Thr remains to be determined. Perhaps the most intriguing group is formed by shorter uncharacterized sequences (small LanBs; red in Fig. 3a) that lack the elimination domain altogether and for which no open reading frame encoding a stand-alone elimination protein is found within nearby genes. These small LanBs are often part of non-ribosomal peptide synthesis clusters (yellow in Fig. 3a), and the observation that they lack the elimination domain indicates that they may add amino acids to a growing peptide in a tRNA-dependent manner.

Further phylogenetic analysis also revealed distinct clades for the elimination domain (Fig. 3b). Some of the enzymes that have been connected to a putative Diels–Alder reaction involved in the biosynthesis of thiopeptides<sup>24</sup> contain a domain that belongs to the same protein family (Pfam) as the LanB elimination domain (PF14028, also called SpaB\_C) and clustered together as a separate clade (purple in Fig. 3b). Hence, despite their limited sequence similarity to the LanB elimination domain (Extended Data Fig. 6), these putative Diels–Alderses will probably contain a similar fold to that presented here.

Our observations predict that aminoacylated tRNA may find much broader participation in natural product biosynthesis than is currently realized<sup>25–28</sup>. The demonstrated use of glutamylated tRNA by NisB to transiently affect the glutamylation of Ser/Thr side chains in NisA is an addition to a growing number of uses of aminoacylated tRNA for functions other than ribosomal protein synthesis<sup>29,30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 March; accepted 23 September 2014.

Published online 26 October 2014.

1. Delves-Broughton, J., Blackburn, P., Evans, R. J. & Hugenoltz, J. Applications of the bacteriocin, nisin. *Antonie van Leeuwenhoek* **69**, 193–202 (1996).
2. Sen, A. K. *et al.* Post-translational modification of nisin. The involvement of NisB in the dehydration process. *Eur. J. Biochem.* **261**, 524–532 (1999).
3. Garg, N., Salazar-Ocampo, L. M. & van der Donk, W. A. *In vitro* activity of the nisin dehydratase NisB. *Proc. Natl Acad. Sci. USA* **110**, 7258–7263 (2013).
4. Lubelski, J., Rink, R., Khusainov, R., Moll, G. N. & Kuipers, O. P. Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cell. Mol. Life Sci.* **65**, 455–476 (2008).
5. Breukink, E. *et al.* Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science* **286**, 2361–2364 (1999).
6. Wiedemann, I. *et al.* Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *J. Biol. Chem.* **276**, 1772–1779 (2001).
7. Hasper, H. E. *et al.* A new mechanism of antibiotic action. *Science* **313**, 1636–1637 (2006).
8. Knerr, P. J. & van der Donk, W. A. Discovery, biosynthesis, and engineering of lantipeptides. *Annu. Rev. Biochem.* **81**, 479–505 (2012).
9. Kaletta, C. & Entian, K. D. Nisin, a peptide antibiotic: cloning and sequencing of the *nisA* gene and posttranslational processing of its peptide product. *J. Bacteriol.* **171**, 1597–1601 (1989).
10. Li, B. *et al.* Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis. *Science* **311**, 1464–1467 (2006).
11. van der Meer, J. R. *et al.* Characterization of the *Lactococcus lactis* nisin A operon genes *nisp*, encoding a subtilisin-like serine protease involved in precursor processing, and *nisR*, encoding a regulatory protein involved in nisin biosynthesis. *J. Bacteriol.* **175**, 2578–2588 (1993).

12. Schnell, N. *et al.* Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature* **333**, 276–278 (1988).
13. Li, C. & Kelly, W. L. Recent advances in thiopeptide antibiotic biosynthesis. *Nat. Prod. Rep.* **27**, 153–164 (2010).
14. Shazman, S. & Mandel-Gutfreund, Y. Classifying RNA-binding proteins based on electrostatic properties. *PLoS Comput. Biol.* **4**, e1000146 (2008).
15. Koehnke, J. *et al.* The cyanobactin heterocyclase enzyme: a processive adenylase that operates with a defined order of reaction. *Angew. Chem. Int. Ed.* **52**, 13991–13996 (2013).
16. Sardar, D., Pierce, E., McIntosh, J. A. & Schmidt, E. W. Recognition sequences and substrate evolution in cyanobactin biosynthesis. *ACS Synth. Biol.* <http://dx.doi.org/10.1021/sb500019b> (13 March 2014).
17. Marques, J. C. *et al.* Processing the interspecies quorum-sensing signal autoinducer-2 (AI-2): characterization of phospho-(S)-4,5-dihydroxy-2,3-pentanedione isomerization by LsrG protein. *J. Biol. Chem.* **286**, 18331–18343 (2011).
18. Mavaro, A. *et al.* Substrate recognition and specificity of NisB, the lantibiotic dehydratase involved in nisin biosynthesis. *J. Biol. Chem.* **286**, 30552–30560 (2011).
19. Plat, A., Kluskens, L. D., Kuipers, A., Rink, R. & Moll, G. N. Requirements of the engineered leader peptide of nisin for inducing modification, export, and cleavage. *Appl. Environ. Microbiol.* **77**, 604–611 (2011).
20. Khusainov, R., Heils, R., Lubelski, J., Moll, G. N. & Kuipers, O. P. Determining sites of interaction between prenisin and its modification enzymes NisB and NisC. *Mol. Microbiol.* **82**, 706–718 (2011).
21. Lubelski, J., Khusainov, R. & Kuipers, O. P. Directionality and coordination of dehydration and ring formation during biosynthesis of the lantibiotic nisin. *J. Biol. Chem.* **284**, 25962–25972 (2009).
22. Zhang, Q. *et al.* Structural investigation of ribosomally synthesized natural products by hypothetical structure enumeration and evaluation using tandem MS. *Proc. Natl Acad. Sci. USA* **111**, 12031–12036 (2014).
23. Lubelski, J., Overkamp, W., Kluskens, L. D., Moll, G. N. & Kuipers, O. P. Influence of shifting positions of Ser, Thr, and Cys residues in prenisin on the efficiency of modification reactions and on the antimicrobial activities of the modified prepeptides. *Appl. Environ. Microbiol.* **74**, 4680–4685 (2008).
24. Li, J. *et al.* ThioFinder: a web-based tool for the identification of thiopeptide gene clusters in DNA sequences. *PLoS ONE* **7**, e45878 (2012).
25. Garg, R. P., Qian, X. L., Alemany, L. B., Moran, S. & Parry, R. J. Investigations of valanimycin biosynthesis: elucidation of the role of seryl-tRNA. *Proc. Natl Acad. Sci. USA* **105**, 6543–6547 (2008).
26. Zhang, W., Ntai, I., Kelleher, N. L. & Walsh, C. T. tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics. *Proc. Natl Acad. Sci. USA* **108**, 12249–12253 (2011).
27. Bougioukou, D. J., Mukherjee, S. & van der Donk, W. A. Revisiting the biosynthesis of dehydrosphingosine reveals a tRNA-dependent pathway. *Proc. Natl Acad. Sci. USA* **110**, 10952–10957 (2013).
28. Gondry, M. *et al.* Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. *Nature Chem. Biol.* **5**, 414–420 (2009).
29. Francklyn, C. S. & Minajigi, A. tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Lett.* **584**, 366–375 (2010).
30. Phizicky, E. M. & Hopper, A. K. tRNA biology charges to the front. *Genes Dev.* **24**, 1832–1860 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank K. Brister and colleagues for facilitating data collection at LS-CAT (Argonne National Labs, Illinois). This work was supported by grants from the National Institutes of Health (NIH) (R01 GM 058822 to W.A.v.d.D., and R01 GM079038 to S.K.N.). M.A.O. was supported partly by a National Institute of General Medical Sciences (NIGMS)–NIH Chemistry–Biology Interface Training Grant (5T32-GM070421) and by the Ford Foundation. Y.H. was supported partly by a Lowell P. Hager fellowship from the Department of Biochemistry. The Bruker UltrafleXtreme MALDI-TOF/TOF mass spectrometer was purchased in part with a grant from the NIH (S10 RR027109 A). The content of this work is solely the responsibility of the authors and do not necessarily represent the official views of the NIH or Ford Foundation.

**Author Contributions** M.A.O. performed all biochemical assays, which were designed and analysed by M.A.O. and W.A.V. Y.H. and S.K.N. performed and interpreted all structural studies. Q.Z. and M.C.W. performed bioinformatic analysis. M.A.O., S.K.N. and W.A.V. wrote the manuscript. M.A.O. and Y.H. contributed equally to this study.

**Author Information** Atomic coordinates and structure factors for the reported crystal structure are deposited in the Protein Data Bank under accession code 4WD9. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.K.N. (s-nair@life.uiuc.edu) or W.A.V. (vddonk@illinois.edu).

## METHODS

Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, Iowa) (Extended Data Table 2). Reagents used for molecular biology were purchased from New England BioLabs (Ipswich, Massachusetts), Thermo Fisher Scientific (Waltham, Massachusetts), or Gold Biotechnology Inc. (St Louis, Missouri). Chemicals were purchased from Sigma-Aldrich (St Louis, Missouri). L-[U-<sup>14</sup>C] glutamic acid (0.1 µCi µl<sup>-1</sup>, 260 mCi mmol<sup>-1</sup>) was purchased from PerkinElmer (Waltham, Massachusetts). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. Cloning inserts were sequenced at ACGT Inc. (Wheeling, Illinois). MALDI-TOF MS analysis was performed with a Bruker UltraflexXtreme MALDI-TOF/TOF (Billerica, Massachusetts) in reflective mode at the University of Illinois Mass Spectrometry Facility. Mass spectrometry data were analysed, smoothed and baseline-corrected with the Bruker Daltonics Software flexAnalysis package. Data were exported, normalized and plotted with Microsoft Excel. Reflective mode was used to resolve mass peaks differing by 18 Da within the 7–8-kDa size range. The small shoulder peaks appearing to the left of unmodified NisA (M) correspond to the laser-induced deamination of the parent peak. Deamination occurs as a consequence of analysing large peptides in reflective mode (Extended Data Fig. 6). All biochemical assays were performed in duplicate with hexahistidine (His<sub>6</sub>)-tagged substrates and enzymes.

**Expression and purification of His<sub>6</sub>-NisA, glutamylated His<sub>6</sub>-NisA and His<sub>6</sub>-NisB.** The *nisA* gene was previously cloned into *pET15b*, resulting in a peptide with an N-terminal His<sub>6</sub> tag<sup>31</sup>. *E. coli* BL21 (DE3) cells (50 µl) were electroporated with the plasmid *his<sub>6</sub>-nisA-pET15b* (50 ng), cells were plated on Luria-Bertani (LB) agar plates supplemented with ampicillin (100 µg ml<sup>-1</sup>) and grown at 37 °C for 12–15 h. A single colony was used to inoculate 50 ml of LB broth supplemented with ampicillin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 4 l of terrific broth (TB) medium supplemented with ampicillin, to an attenuation at 600 nm (*D*<sub>600</sub>) of 0.025. Cultures were grown at 37 °C to a final *D*<sub>600</sub> of 1.0. Peptide expression was induced by the addition of IPTG (isopropyl β-D-thiogalactoside) to a final concentration of 1 mM, and cultures were grown at 18 °C for 18 h. Peptide purification was performed by following a previously described method<sup>31</sup>. The cloning, expression and purification of His<sub>6</sub>-NisB and glutamylated His<sub>6</sub>-NisA are described elsewhere<sup>3</sup>.

**Preparation of *E. coli* BL21 (DE3) cell extract.** The cell extract of *E. coli* BL21 (DE3) was prepared by following a previously described method<sup>32</sup>, with minor modifications. *E. coli* BL21 (DE3) cells were inoculated in 2 l of TB medium and grown at 37 °C to a final *D*<sub>600</sub> of 1.0. Cells were harvested (4,650g, 10 min, 4 °C) and the pellet was washed three times by resuspension in S30 buffer (10 mM Tris-acetate buffer pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate, 1 mM dithiothreitol (DTT), 0.3 mM EDTA, 0.3 mM MgCl<sub>2</sub>) followed by centrifugation (4,650g, 10 min, 4 °C). Cells were then resuspended in 1 ml of S30 buffer per gram of wet cells and lysed with a MultiFlex C3 Homogenizer (Avestin). The cell lysate was centrifuged twice (22,789g, 30 min, 4 °C) and the supernatant was dialysed four times against 50 volumes of S30 buffer (without DTT) using a dialysis cassette with a molecular mass cutoff (MWCO) of 10 kDa. The cell extract was then centrifuged (4,000g, 10 min, 4 °C) and the supernatant was frozen and stored in 1-ml samples at –80 °C for future use.

**Anion-exchange fractionation of *E. coli* cell extract.** A 1-ml sample of cell extract was loaded onto a 1 ml Hi Trap DEAE FF (GE Healthcare) column equilibrated with five column volumes of start buffer (10 mM Tris-HCl pH 8.0), using an AKTA FPLC system (Amersham Pharmacia Biosystems). The column was washed with ten column volumes of start buffer and the sample was eluted with a gradient of 0–100% (v/v) of elution buffer (10 mM Tris-HCl, 1 M NaCl pH 8.0) in start buffer over 20 column volumes. Eluent was detected by absorbance at 280 nm (Extended Data Fig. 1). Collected fractions were desalted by washing five times with 500 µl of S30 buffer (without DTT), using a 10-kDa MWCO Amicon centrifugal filter (Millipore) (14,000g, 10 min, 4 °C). Fractions were stored at –80 °C.

**Cloning, expression and purification of *E. coli* BL21 (DE3) GluRS.** The gene *gluX* (WP\_024237069) was cloned into *pET28a*, resulting in a protein with an N-terminal His<sub>6</sub> tag. Genomic DNA (gDNA) from *E. coli* BL21 (DE3) cells was extracted with an UltraClean Microbial DNA isolation kit (MoBio) in accordance with the manufacturer's protocol. The *gluX* gene was PCR amplified in the following reaction mixture: Phusion HF buffer (1×), dNTP (0.2 mM), primers (1 µM each; Extended Data Table 2), gDNA template (10 ng), dimethylsulphoxide (3% v/v) and Phusion polymerase (0.04 U µl<sup>-1</sup>) in a total volume of 50 µl. The amplification was performed in 20 cycles: denaturation (98 °C for 60 s), annealing (67 °C for 60 s) and amplification (72 °C for 42 s). The PCR product was purified by gel extraction on a 1% (w/v) agarose gel using the QIAquick Gel Extraction Kit (Qiagen). The vector *pET28a* and *gluX* PCR fragment were digested with NdeI and XhoI (NEB) restriction endonucleases and purified by gel extraction as described above; the resulting fragments were ligated in 20 µl using T4 DNA ligase (NEB) at 16 °C for

18 h. A 10-µl sample from the ligation reaction was used to transform *E. coli* DH5α (DE3) cells using the heat shock method. The cells were plated on LB plates supplemented with kanamycin (50 µg ml<sup>-1</sup>), and the plates were incubated at 37 °C for 12–15 h. Single colonies were picked and grown in LB supplemented with kanamycin at 37 °C for 12–15 h, and the plasmid *his<sub>6</sub>-gluX-pET28a* was isolated with a QIA prep Spin Miniprep Kit (Qiagen). Insert integrity was verified by sequencing the plasmid with the appropriate primers.

*E. coli* BL21 (DE3) cells (50 µl) were electroporated with *his<sub>6</sub>-gluX-pET28a* (50 ng), plated on LB agar plates supplemented with kanamycin, and grown at 37 °C for 12–15 h. A single colony was used to inoculate 50 ml of LB broth supplemented with kanamycin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 3 l of TB medium supplemented with kanamycin to a *D*<sub>600</sub> of 0.025. Cultures were grown at 37 °C to a *D*<sub>600</sub> of 0.5. Protein expression was induced by the addition of IPTG to a final concentration of 0.2 mM, and cultures were grown at 18 °C for 20 h. Protein purification was performed as described for His<sub>6</sub>-NisB<sup>3</sup>. After elution from a 5-ml HisTrap HP column (GE Healthcare), the protein was desalted with a PD-10 desalting column (GE Healthcare), concentrated with a 30-kDa MWCO Amicon centrifugal filter (Millipore) (4,000g, 20 min, 4 °C) and stored in storage buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol) at –80 °C.

**Cloning, expression and purification of *L. lactis* HP GluRS.** The gene *gluX* (WP\_011836000) from *L. lactis* HP was cloned into *pET28a*, resulting in a protein with an N-terminal His<sub>6</sub> tag. gDNA from *L. lactis* HP cells was extracted with an UltraClean Microbial DNA isolation kit (MoBio) in accordance with the manufacturer's protocol. The *gluX* gene was amplified by PCR, under similar conditions to those described above, with appropriate primers (Extended Data Table 2). The amplification was performed in 20 cycles: denaturation (98 °C for 60 s), annealing (65 °C for 60 s) and amplification (72 °C for 60 s). The PCR purification, ligation, expression and purification of *L. lactis* GluRS were performed as described above for *E. coli* GluRS.

**Purification of *E. coli* tRNA<sup>Glu</sup> and glutamyl-tRNA<sup>Glu</sup>.** Primers for *E. coli* tRNA<sup>Glu</sup> were designed in accordance with a method described previously<sup>33</sup> (Extended Data Table 2). The tRNA<sup>Glu</sup> double-stranded DNA template for *in vitro* transcription was prepared by filling 5' overhangs in the following reaction mixture: NEB Buffer 2 (1×), primers (4 µM each), dNTP (33 µM), NEB DNA pol I (large Klenow fragment) (1 U µg<sup>-1</sup> DNA), in a final volume of 50 µl. The reaction was incubated at 25 °C for 15 min, quenched with EDTA (10 mM) and incubated at 75 °C for 25 min; double-stranded DNA tRNA<sup>Glu</sup> template was precipitated with ethanol. *In vitro* transcription was performed with a method described previously<sup>34</sup>.

Glutamyl-tRNA<sup>Glu</sup> was performed in the following reaction mixture: tRNA<sup>Glu</sup> (50 µM), L-glutamate (10 mM), ATP (10 mM), *E. coli* GluRS (5 µM), thermostable inorganic pyrophosphatase (TIPP) (New England BioLabs) (0.04 U µl<sup>-1</sup>), HEPES pH 7.5 (100 mM), KCl (10 mM), MgCl<sub>2</sub> (20 mM), DTT (1 mM), in a final volume of 500 µl. The reaction was incubated at 37 °C for 1 h and desalted by washing ten times with 500 µl HEPES pH 7.5 (100 mM) using a 10-kDa MWCO Amicon centrifugal filter (Millipore) (14,000g, 10 min, 4 °C). Glutamyl-tRNA<sup>Glu</sup> was then purified by acidic phenol extraction with a method described previously<sup>35</sup>.

To assess the extent of aminoacylation on *E. coli* tRNA<sup>Glu</sup> transcribed *in vitro*, glutamyl-tRNA<sup>Glu</sup> was analysed by denaturing acidic polyacrylamide gel with a procedure described previously<sup>35,36</sup>. tRNA aminoacylation assays were performed with L-[U-<sup>14</sup>C]glutamic acid (0.1 µCi µl<sup>-1</sup>, 260 mCi mmol<sup>-1</sup>) in the following reaction mixture: *in vitro* transcribed *E. coli* tRNA<sup>Glu</sup> (3 µM), L-[U-<sup>14</sup>C]glutamic acid (50 µM), HEPES pH 7.5 (100 mM), ATP (5 mM), tris(2-carboxyethyl)phosphine (TCEP) (1 mM), KCl (10 mM), MgCl<sub>2</sub> (20 mM), TIPP (0.02 U µl<sup>-1</sup>), purified *E. coli* GluRS (1 µM), in a final volume of 50 µl. The reaction was incubated at 37 °C for 1 h and quenched with sodium acetate pH 5.6 (300 mM). Glutamyl-tRNA<sup>Glu</sup> was precipitated by standard precipitation with ethanol. The resulting pellet was dissolved in 10 µl of acidic loading buffer (100 mM sodium acetate pH 5.6, 7 M urea, 0.05% bromophenol blue) and the sample was loaded on a 7% acidic polyacrylamide gel (100 mM sodium acetate, 8 M urea). Electrophoresis was performed for 120 min at 100 V (constant voltage) in an ice bath, using 100 mM sodium acetate pH 5.6 as running buffer. To visualize tRNA, the gel was stained with staining solution (500 mM sodium acetate pH 5.6, 0.06% methylene blue) for 30 min and destained in water. After drying the gel overnight with a gel-drying kit (Promega) in accordance with the manufacturer's procedure, glutamyl-tRNA<sup>Glu</sup> was visualized by exposing the dried gel on a phosphorimager cassette for 4 days. The film was then visualized with a STORM 840 Phosphorimager Scanner (Amersham Biosciences) (Extended Data Fig. 2).

**RNA extraction from *L. lactis* HP.** *L. lactis* HP cells were grown for 16 h in 2 l of M17 medium supplemented with 0.5% (w/v) glucose at 30 °C. Cells were washed three times with S30 buffer, and cell extract was prepared as described above for *E. coli* cells. RNA was extracted from the cell extract with acidic phenol. In brief, a 500-µl sample of cell extract was mixed with 500 µl of acidic phenol (pH ~4.5) and the resulting emulsion was vortex-mixed for 1 min. The sample was centrifuged



(14,000g, 5 min, 4 °C) and the aqueous phase was collected for a second extraction with acidic phenol. After centrifugation, the aqueous layer was extracted twice with 500 µl of chloroform:isoamyl alcohol (24:1). The sample was vortex-mixed for 1 min and centrifuged (14,000g, 5 min, 4 °C). Total RNA was precipitated by standard precipitation with ethanol and dissolved in RNase-free distilled water. To remove any aminoacylated tRNA, total RNA (0.66 µg µl<sup>-1</sup>) was incubated for 1 h at 37 °C in EDTA (0.6 mM) and Tris-HCl pH 9.0 (66 mM). The sample was acidified with sodium acetate (300 mM) and RNA was extracted by standard precipitation with ethanol<sup>36</sup>.

**In vitro dehydration assays.** The following reaction mixture was used for dehydration assays: *E. coli* cell extract/fraction (10 µl), HEPES pH 7.5 (100 mM), DTT (1 mM), L-glutamate (10 mM), NisA (200 µM), MgCl<sub>2</sub> (10 mM), KCl (10 mM), NisB (5 µM), TIPP (0.02 U µl<sup>-1</sup>), ATP (5 mM), in a final volume of 30 µl. The assay was incubated at 30 °C for 5 h, centrifuged to remove insoluble material (14,000g, 5 min, 25 °C), and desalted with a C-18 ZipTip concentrator (Millipore). The sample was mixed in a 1:1 ratio with sinapinic acid matrix and spotted on a Bruker MALDI-TOF plate. When noted, the cell extract (10 µl) was replaced with *E. coli* GluRS (5 µM) and *E. coli* tRNA<sup>Glu</sup> (50 µM). In addition, when noted, the cell extract (10 µl) was treated with either RNase (10 U) or RNase-free DNase (10 U) in the presence of CaCl<sub>2</sub> (100 µM), and incubated at 25 °C for 1 h before dehydration assays.

For dehydration assays using glutamyl-tRNA<sup>Glu</sup> the following reaction mixture was used: HEPES pH 7.5 (100 mM), DTT (500 µM), MgCl<sub>2</sub> (10 mM), KCl (10 mM), His<sub>6</sub>-NisA (10 µM), Glu-tRNA<sup>Glu</sup> (~100 µM), His<sub>6</sub>-NisB (5 µM), in a final volume of 30 µl. The assay was incubated at 30 °C for 12 h and prepared as described above for analysis by MALDI-TOF MS. In a separate experiment, ATP (1 mM) was added to the reaction mixture.

For dehydration assays using *L. lactis* HP total RNA the following reaction mixture was used: *L. lactis* RNA (1 µg µl<sup>-1</sup>), glutamate (10 mM), MgCl<sub>2</sub> (10 mM), KCl (10 mM), HEPES pH 7.5 (50 mM), *L. lactis* GluRS (10 µM), NisB (5 µM), NisA (100 µM), ATP (5 mM), TIPP (0.02 U µl<sup>-1</sup>), TCEP (1 mM). The assay was incubated at 30 °C for 3 h and prepared as described above for analysis by MALDI-TOF MS.

**Glutamate elimination assays.** The glutamylated precursor peptide NisA was incubated with endoprotease ArgC in the following reaction mixture: glutamylated NisA (500 µM), Tris-HCl (20 mM), CaCl<sub>2</sub> (10 mM), EDTA (0.5 mM), DTT (5 mM), ArgC (0.005 µg µl<sup>-1</sup>). The reaction was incubated for 4 h at 37 °C. Glutamylated core and leader peptide were purified by analytical high-performance liquid chromatography (HPLC) on a Phenomenex C<sub>18</sub> column (Luna; 250 mm × 4.60 mm, particle size 10 µm, pore size 100 Å) connected to an Agilent 1260 infinity liquid chromatography system. Peptides were separated by using a linear gradient of 2% (v/v) solvent B (80% (v/v) acetonitrile, 20% (v/v) H<sub>2</sub>O, 0.086% (v/v) trifluoroacetic acid) in solvent A (0.1% (v/v) trifluoroacetic acid in water) to 100% (v/v) solvent B over 60 min at a flow rate of 1 ml min<sup>-1</sup>. Peptide elution was monitored using a 220-nm wavelength. Desired fractions were freeze-dried and stored at -20 °C.

The following reaction mixture was used for the glutamate elimination assays: HEPES pH 7.5 (100 mM), glutamylated His<sub>6</sub>-NisA core peptide (~100 µM), NisB (5 µM), MgCl<sub>2</sub> (10 mM), KCl (10 mM), TCEP (1 mM). The reaction was performed in a 15 µl volume, incubated for 60 min at 30 °C and analysed by MALDI-TOF MS as described above.

**Crystallization and structure determination of NisB.** His<sub>6</sub>-tagged NisB was co-expressed with untagged NisA and purified as described previously<sup>3</sup>. Before crystallization, samples of NisB were methylated with a method described previously<sup>37</sup> and purified again by size-exclusion chromatography (HiLoad 16/600 Superdex 75 pg column; GE Healthcare) into a buffer consisting of HEPES pH 7.5 (20 mM) and KCl (300 mM). Crystals of methylated NisB (5 mg ml<sup>-1</sup>) were grown at 9 °C with the vapour diffusion method, using a precipitant of 100 mM Bicine (pH 8.5) and 15–25% poly(ethylene glycol) 6000. Macro-seeding and micro-seeding facilitated the formation of crystals suitable for diffraction data collection. SeMet NisB was expressed, purified and crystallized in a similar manner. Crystallographic phases were determined by single-wavelength anomalous diffraction methods from data collected on crystals of SeMet NisB to a resolution limit of 3.2 Å. Native and SeMet data were collected at Sector 21 ID (LS-CAT, Advanced Photon Source, Argonne National Labs, Illinois) and data were integrated and scaled using HKL2000 (ref. 38) or XDS<sup>39</sup>. Heavy-atom sites were located using the SHELX suite of programs<sup>40</sup>, and refinement of heavy-atom parameters in SHARP<sup>41</sup> yielded an initial figure of merit of 0.276. Density modification and two-fold symmetry averaging produced experimental maps that permitted most secondary structural elements to be manually assigned with COOT<sup>42</sup>. Further model building used the 2.9-Å resolution data collected on crystals of native NisB. Although obvious electron density corresponding to the NisA peptide (co-purified from the coexpression with NisB) can be observed in experimentally phased maps, the peptide was only built into the model once the free *R* factor for the structure had decreased to below 0.30. The validity of all models

was routinely determined with MOLPROBITY and by using the free *R* factor to monitor improvements during building and crystallographic refinement.

**Coexpression of NisB-R784A with NisA.** Site-directed mutagenesis was performed on *his<sub>6</sub>-nisA/nisBpRSF Duet-1* (ref. 3) to replace the conserved Arg 784 in NisB with Ala, using a method described previously<sup>43</sup> with appropriate primers (Extended Data Table 2). The resultant plasmid was used for coexpression of His<sub>6</sub>-NisA with NisB-R784A, and modified His<sub>6</sub>-NisA was purified in accordance with a method described previously<sup>3</sup>.

**Construction of the LanB protein sequence similarity map.** The LanB sequences were obtained by iterative Psi-Blast search, and the hits were grouped into full-length LanB and small LanB based on whether they had the SpaB-C-terminal domain. Thiopeptide dehydratase sequences were obtained from the ThioBase database<sup>24</sup>. Network analysis was performed by BLASTP searches, comparing each sequence against each other. A Matlab script was written to remove all duplicate comparisons, and the result was imported into the Cytoscape software package. Each node in the network represents a protein sequence, and each edge represents the pairwise connection between two sequences with a BLASTP *E* value lower than the cut-off value. The nodes were arranged by using the yFiles organic layout provided with Cytoscape version 2.8.3. Accession numbers are listed in Supplementary Table 1.

**Construction of the phylogenetic tree.** Sequences for the SpaB C-terminal-domain-containing proteins were obtained by searching the NCBI assembled bacterial genomes database with the SpaB\_C PFAM<sup>44</sup> hidden Markov model, using HMMER3 (ref. 45). Putative Diels-Alderase sequences were identified constructing a hidden Markov model from an alignment of NosO orthologues obtained from the ThioBase<sup>24</sup> database. This hidden Markov model was used in the same manner as the SpaB\_C hidden Markov model. The portions of protein sequences that aligned to the SpaB\_C PFAM or the putative Diels-Alderase hidden Markov model were combined and clustered with CD-HIT<sup>46</sup>, identifying representative sequences that shared no more than 50% identity. The genomic contexts of these representative sequences were examined to determine the type of protein and biosynthetic cluster in which the domains resided. Then, the sequences were aligned with PRALINE<sup>47</sup>. The alignment was then adjusted manually to remove poorly aligning regions and subsequently used to construct a phylogenetic tree using MrBayes<sup>48</sup>. Likelihood values were calculated over 7,500,000 generations of two runs with four chains each. For the analysis, among-site rate variation was considered to be invariable for a portion of sites and the rates for the other sites were drawn from a gamma distribution estimated with four rate categories. Furthermore, the WAG amino acid substitution model and random starting trees were utilized.

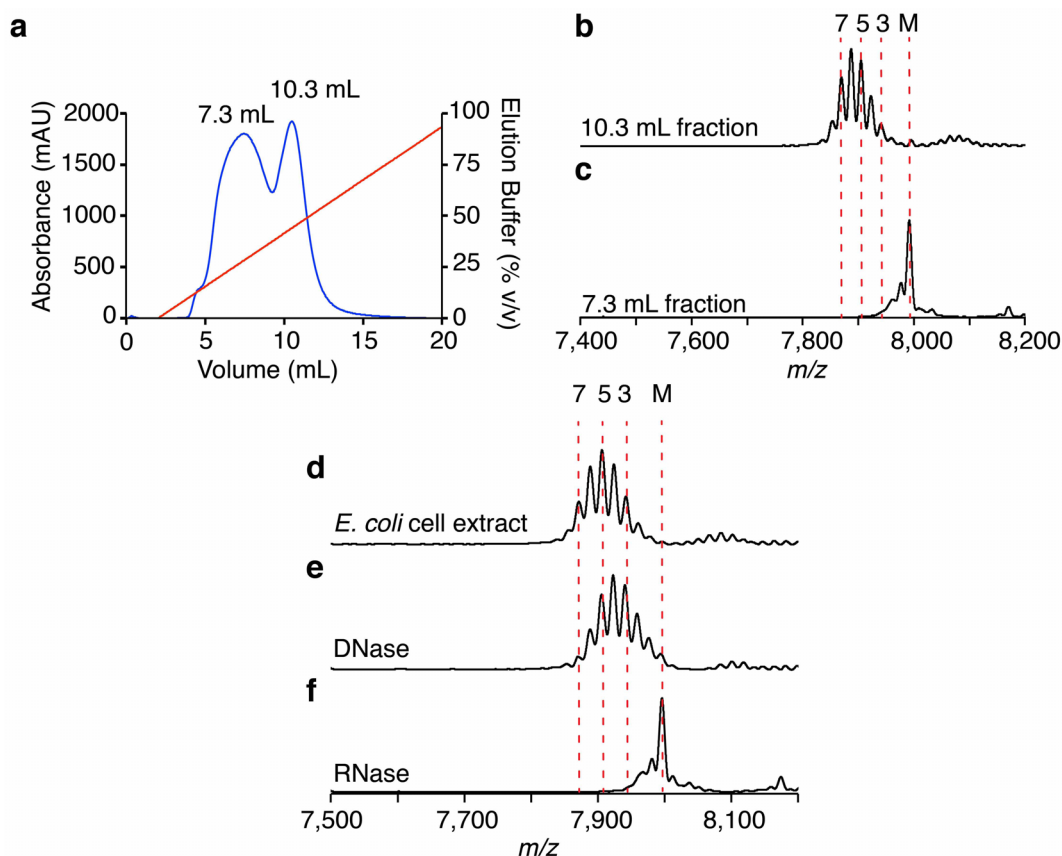
The final tree was generated with a 25% relative burn-in. Accession numbers are listed in Supplementary Table 1.

**Representative sequence alignment.** Representative sequences for thiopeptide dehydratases and putative Diels-Alderas were selected from ThioBase<sup>24</sup>. The SpaB C-terminal domains of these proteins were identified by HMMER as described above and aligned with the SpaB C-terminal domains of characterized full-length LanB proteins, using MUSCLE<sup>49</sup>.

- Li, B., Cooper, L. E. & van der Donk, W. A. *In vitro* studies of lantibiotic biosynthesis. *Methods Enzymol.* **458**, 533–558 (2009).
- Kigawa, T. *et al.* Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J. Struct. Funct. Genomics* **5**, 63–68 (2004).
- Sherlin, L. D. *et al.* Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA* **7**, 1671–1678 (2001).
- Rio, D. C., Ares, M. J., Hannon, G. J. & Nilsen, T. W. *RNA: A Laboratory Manual* 216–219 (Cold Spring Harbor Laboratory Press, 2011).
- Walker, S. E. & Fredrick, K. Preparation and evaluation of acylated tRNAs. *Methods* **44**, 81–86 (2008).
- Janssen, B. D., Diner, E. J. & Hayes, C. S. Analysis of aminoacyl- and peptidyl-tRNAs by gel electrophoresis. *Methods Mol. Biol.* **905**, 291–309 (2012).
- Walter, T. S. *et al.* Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617–1622 (2006).
- Otwiński, Z., Borek, D., Majewski, W. & Minor, W. Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A* **59**, 228–234 (2003).
- Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
- Thorn, A. & Sheldrick, G. M. Extending molecular-replacement solutions with SHELXE. *Acta Crystallogr. D* **69**, 2251–2256 (2013).
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59**, 2023–2030 (2003).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
- Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).
- Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).



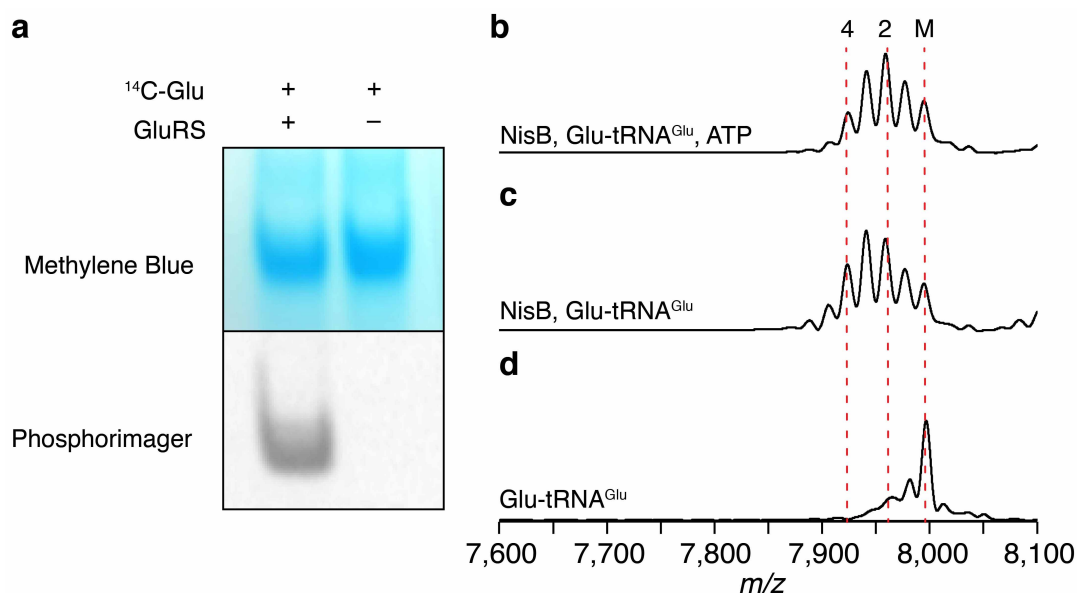
46. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
47. Simossis, V. A. & Heringa, J. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res.* **33**, W289–W294 (2005).
48. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
49. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).



#### Extended Data Figure 1 | RNA-dependent dehydration of His<sub>6</sub>-NisA.

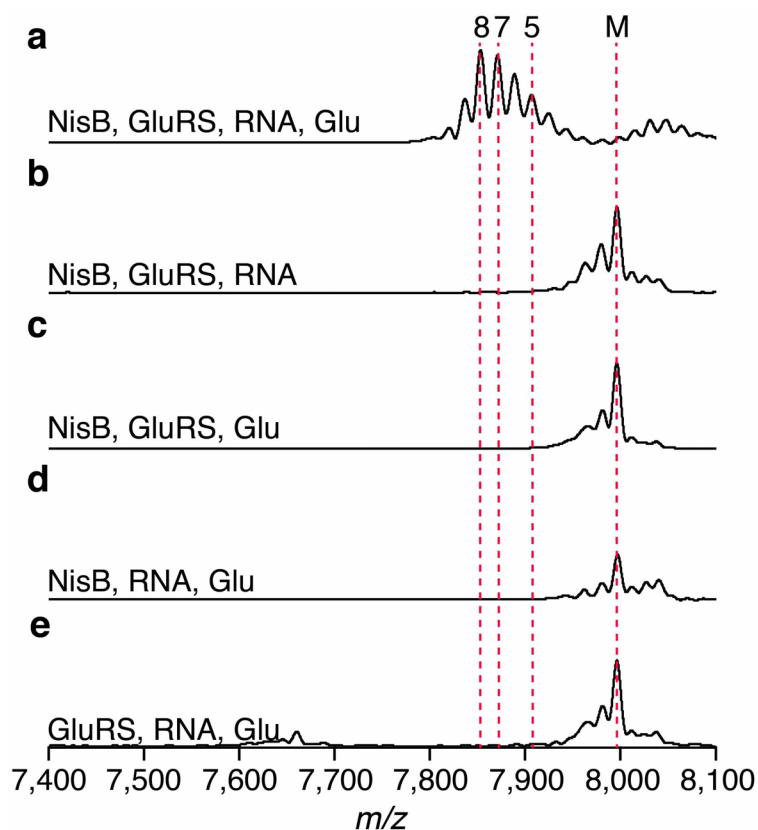
**a**, Anion exchange chromatogram of *E. coli* cell extract. Fractionation was monitored at 280 nm (blue) and the sample was eluted with a NaCl gradient (red). Two peaks were observed at retention volumes of 7.3 and 10.3 ml, respectively. **b–f**, MALDI-TOF mass spectra of His<sub>6</sub>-NisA after *in vitro* reaction with His<sub>6</sub>-NisB, ATP and Glu, in the presence of the 10.3-ml fraction (**b**), the 7.3-ml fraction (**c**), *E. coli* cell extract (**d**), and *E. coli* cell extract

treated with DNase (**e**) or RNase (**f**). Numbers above the peaks correspond to the number of dehydrations of His<sub>6</sub>-NisA after incubation with His<sub>6</sub>-NisB (**b–f**). M, unmodified His<sub>6</sub>-NisA ( $m/z$  7,992, calc.  $m/z$  7,996); 3, threefold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,941, calc.  $m/z$  7,942); 5, fivefold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,905, calc.  $m/z$  7,906); 7, sevenfold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,870, calc.  $m/z$  7,870). Assays were performed in duplicate.



**Extended Data Figure 2 | Glutamyl-tRNA dependent dehydration of His<sub>6</sub>-NisA.** **a**, Glutamylation of *in vitro* transcribed *E. coli* tRNA<sup>Glu</sup> by purified *E. coli* GluRS using L-[U-<sup>14</sup>C]glutamic acid was analysed by gel electrophoresis. The gel was stained with methylene blue (top), dried, exposed to a phosphorimager screen, and scanned (bottom). Thus, recombinant purified GluRS is capable of aminoacylating *E. coli* tRNA<sup>Glu</sup> lacking post-transcriptional modifications. The gel was cropped for visualization purposes. **b–d**, MALDI-TOF mass spectra of the precursor peptide His<sub>6</sub>-NisA

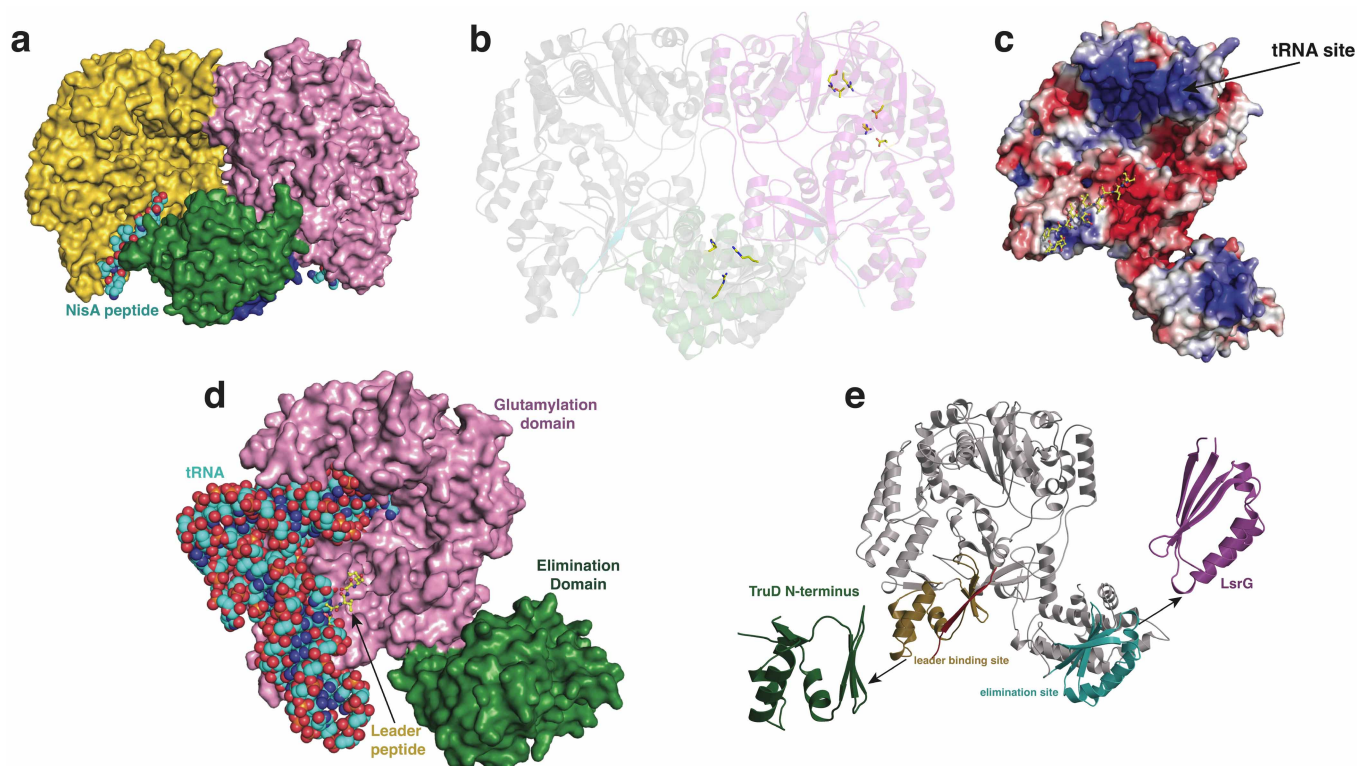
after incubation with His<sub>6</sub>-NisB, glutamyl-tRNA<sup>Glu</sup> and ATP (**b**), His<sub>6</sub>-NisB and glutamyl-tRNA<sup>Glu</sup> (**c**), and glutamyl-tRNA<sup>Glu</sup> in the absence of His<sub>6</sub>-NisB (**d**). Numbers above the mass spectra correspond to the number of dehydrations of His<sub>6</sub>-NisA after incubation with His<sub>6</sub>-NisB. M, unmodified His<sub>6</sub>-NisA ( $m/z$  7,995, calc.  $m/z$  7,996); 2, twofold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,959, calc.  $m/z$  7,960); 4, fourfold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,924, calc.  $m/z$  7,924). Dehydration assays were performed in duplicate.



**Extended Data Figure 3 | His<sub>6</sub>-NisB-catalysed dehydration of His<sub>6</sub>-NisA in the presence of *L. lactis* HP RNA and GluRS.** MALDI-TOF MS analysis of His<sub>6</sub>-NisA after incubation with His<sub>6</sub>-NisB, *L. lactis* His<sub>6</sub>-GluRS, RNA isolated from *L. lactis*, glutamate and ATP (a), His<sub>6</sub>-NisB, *L. lactis* His<sub>6</sub>-GluRS and RNA, and ATP (b), His<sub>6</sub>-NisB, *L. lactis* His<sub>6</sub>-GluRS, Glu and ATP (c), His<sub>6</sub>-NisB, *L. lactis* RNA, Glu and ATP (d), and *L. lactis* His<sub>6</sub>-GluRS and

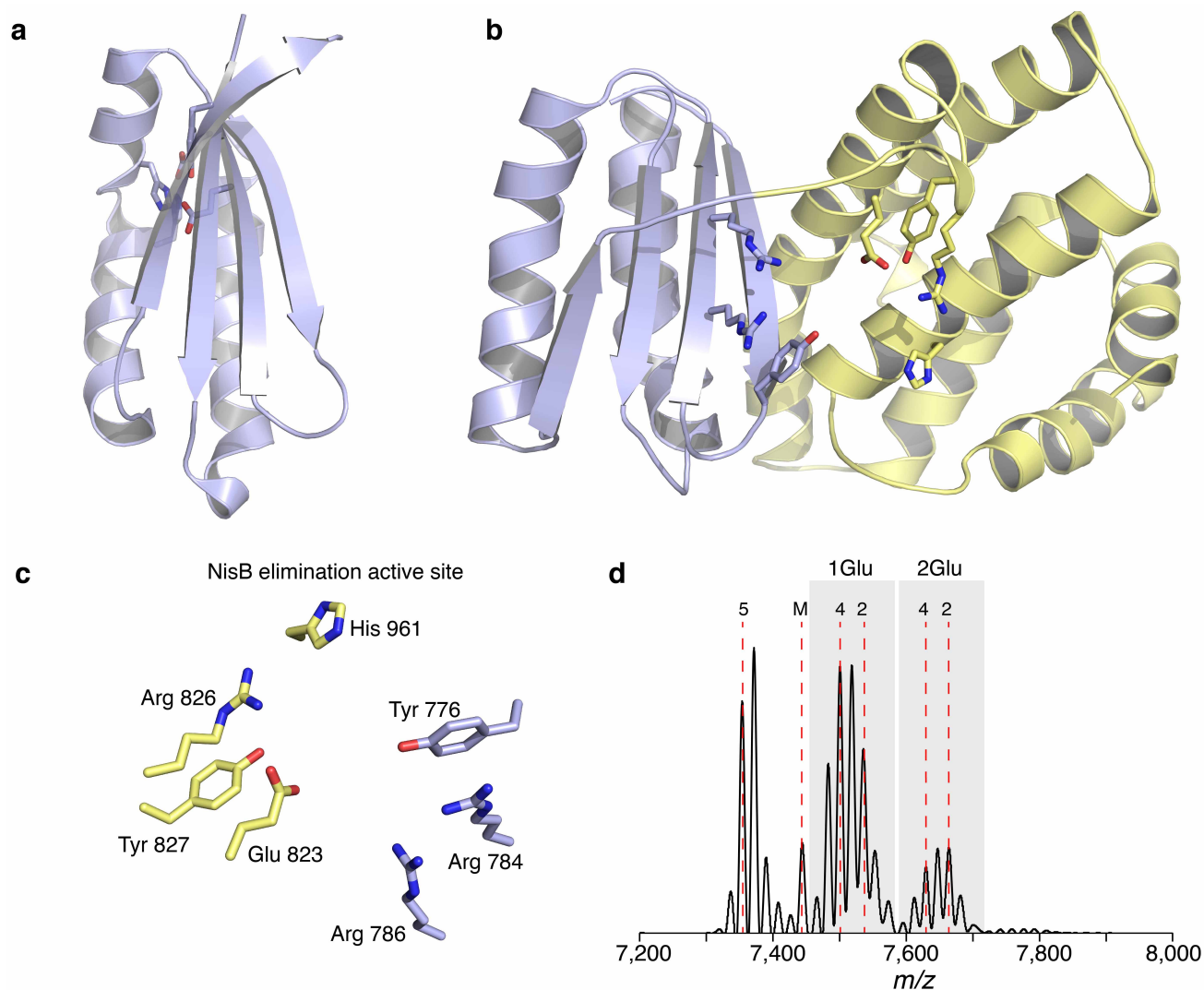
RNA, Glu and ATP (e). Numbers above the mass spectra correspond to the number of dehydrations of His<sub>6</sub>-NisA after incubation with His<sub>6</sub>-NisB. M, unmodified His<sub>6</sub>-NisA ( $m/z$  7,997, calc.  $m/z$  7,996); 5, fivefold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,906, calc.  $m/z$  7,906); 7, sevenfold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,870, calc.  $m/z$  7,870); 8, eightfold dehydrated His<sub>6</sub>-NisA ( $m/z$  7,853, calc.  $m/z$  7,852). Dehydration assays were performed in duplicate.





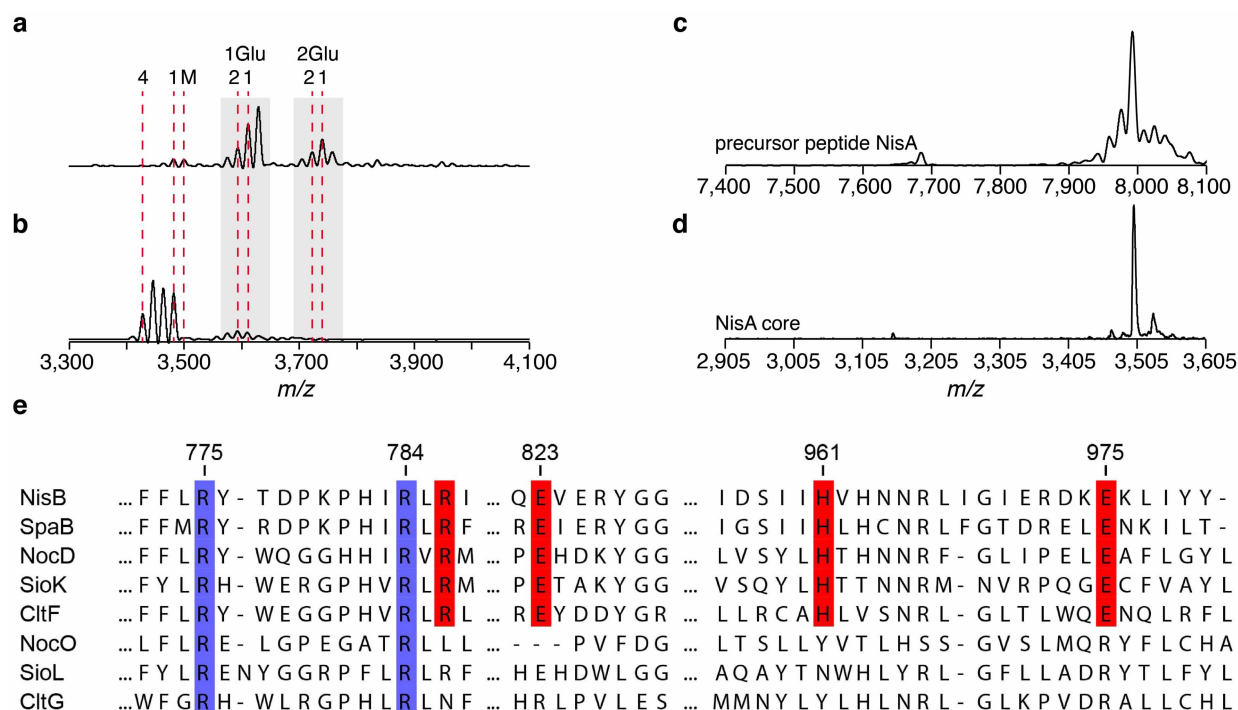
**Extended Data Figure 4 | Surface representation, structural homology, and model for tRNA engagement by NisB.** **a**, The NisB homodimer is shown with one monomer coloured in gold (glutamylation domain) and blue (glutamate elimination domain); the other monomer is coloured pink (glutamylation domain) and green (glutamate elimination domain). The NisA peptide is shown as spheres. **b**, Transparent cartoon representation of the dimer showing the clustering of the residues important for glutamylation (in the pink domain) and glutamate elimination (in the green domain) represented as yellow sticks. **c**, Calculated electrostatic potential mapped onto the NisB surface showing the basic patch (blue) that probably engages the glutamyl-tRNA<sup>Glu</sup>. The NisA peptide is shown in yellow. **d**, tRNA<sup>Glu</sup>-NisB binding model with the

NisB glutamylation domain in pink and the elimination domain in green. The leader peptide is shown in a yellow ball-and-stick representation. The double-stranded RNA-binding protein A complexed with its cognate RNA (PDB 1DI2) was used to derive a NisB docking pose for binding to bacterial tRNA<sup>Glu</sup> (*T. thermophilus* tRNA taken from PDB 1N78). The model results in the placement of the aminoacylated CCA terminus in the vicinity of residues that have been shown to be important for glutamylation activity. **e**, Domains within the overall structure of NisB that share notable homology are shown for the leader-peptide-binding site (gold) and the site for glutamate elimination (cyan). Structures that are related by homology are shown adjacent to the respective domains.



**Extended Data Figure 5 | Structural analysis of the glutamate elimination domain in NisB.** **a**, Diagrammatic representation of the LsrG protein (PDB 3QM). Residues involved in LsrG catalysis are shown as sticks. The putative LsrG active site, located between the  $\alpha$ -helices and the  $\beta$ -sheet, was proposed on the basis of the presence of an unidentified ligand<sup>17</sup>. Mutations of residues in the proposed active site of LsrG demonstrated their importance in activity, but no functions were assigned to individual residues<sup>17</sup>. **b**, Diagrammatic representation of the NisB glutamate elimination domain. The segment with structural homology to LsrG is coloured light blue; the remainder of the elimination domain is coloured yellow. Residues important for glutamate elimination or dehydration<sup>3</sup> are shown as sticks. **c**, Residues

important for glutamate elimination and net dehydration delineate a putative active site for glutamate elimination. Residues Arg 786, Arg 826 and His 961 are important for glutamate elimination<sup>3</sup>, and Glu 823 has previously been shown to be partly important for dehydration<sup>3</sup>. The importance of Arg 784 in the glutamate elimination step was determined in this study. **d**, MALDI-TOF MS analysis of His<sub>86</sub>-NisA purified after coexpression with His<sub>86</sub>-NisB-R784A. The presence of multiple glutamylated intermediates demonstrates that Arg 784 is important for glutamate elimination and not for glutamylation. The designations 1Glu and 2Glu above the peaks indicate the number of glutamate adducts on the family of peaks, with the number shown below indicating the additional number of dehydrations for each peak.



**Extended Data Figure 6 | NisB-catalysed glutamate elimination of glutamylated NisA core peptide, laser-induced deamination of full-length precursor peptide NisA, and sequence analysis of the glutamate elimination domain in NisB.**

**a, b**, MALDI-TOF MS analysis of glutamylated NisA core peptide before (**a**) and after (**b**) incubation with His<sub>6</sub>-NisB. The data show that the leader peptide is not required for NisB-catalysed glutamate elimination.

Numbers above correspond to the number of glutamate adducts or dehydrations of NisA core peptide. M, NisA core peptide ( $m/z$  3,499, calc.  $m/z$  3,498); 1, onefold dehydrated NisA core peptide ( $m/z$  3,481, calc.  $m/z$  3,480); 4, fourfold dehydrated NisA core peptide ( $m/z$  3,427, calc.  $m/z$  3,426); 1Glu-1, NisA core peptide after formation of one glutamate adduct and one dehydration ( $m/z$  3,611, calc.  $m/z$  3,609); 1Glu-2, NisA core peptide after formation of one glutamate adduct and two dehydrations ( $m/z$  3,593, calc.  $m/z$  3,591); 2Glu-1, NisA core peptide after formation of two glutamate adducts and one dehydration ( $m/z$  3,740, calc.  $m/z$  3,738); 2Glu-2, NisA core peptide after formation of two glutamate adducts and two dehydrations ( $m/z$  3,722, calc.  $m/z$

3,720). **c**, MALDI-TOF MS in reflective mode of the precursor peptide NisA.

**d**, MALDI-TOF MS in reflective mode of NisA core peptide after treatment with the protease ArgC. ArgC cleaves after Arg-1 in the leader peptide of NisA (Fig. 1a). MALDI-TOF MS analysis of the precursor peptide NisA in reflective mode caused the appearance of smaller shoulder peaks next to the parent mass. These shoulder peaks correspond to laser-induced deamination of the parent mass and are observed only for high-molecular-mass peptides (>6 kDa). The shoulder peaks were not observed after proteolytic removal of the leader peptide, confirming that they were not the result of dehydrations. **e**, Sequence alignment of selected elimination domains (SpaB\_C PFAM family) present in full LanBs (NisB and SpaB), thiopeptide dehydratases (NocD, SioK and CltF) and putative<sup>24</sup> thiopeptide Diels-Alderase (NocO, SioL and CltG). Residues involved in glutamate elimination (red) are conserved in full LanBs and thiopeptide dehydratases but not in the putative Diels-Alderase.

Extended Data Table 1 | Data collection, phasing and refinement statistics for NisB

	Native 1	SeMet
<b>Data collection</b>		
a, b, c (Å)	98.5, 107.3, 135.7	100.8, 108.5, 136.5
$\beta$ (°)	109.9	110.1
Resolution (Å) *	50.00 -2.9 (3.0 - 2.9)	50.00 - 3.2 (3.2 - 3.26)
Total reflections	257,756	217,912
Unique reflections	58,894	45,467
R <sub>sym</sub> (%)	10.2 (87.1)	7.9 (82.5)
I/ $\sigma$ (I)	10.6 (1.9)	17.9 (2.0)
Completeness (%)	99.6 (99.8)	99.2 (98.7)
Redundancy	4.4 (4.4)	4.8 (4.8)
<b>Phasing</b>		
Figure of Merit <sup>†</sup>		0.276/0.76
<b>Refinement</b>		
Resolution (Å)	34.5 – 2.9	
No. reflections used	58,852	
R <sub>work</sub> /R <sub>free</sub> <sup>‡</sup>	19.4/25.7	
<b>Number of atoms</b>		
Protein	15,901	
Peptide	200	
<b>B-factors</b>		
Protein	27.3	
Peptide	38.6	
<b>R.m.s deviations</b>		
Bond lengths (Å)	0.011	
Bond angles (°)	1.36	

\* Highest resolution shell is shown in parenthesis.

<sup>†</sup> Figure of merit- Probability weighted average of the cosine of the phase error, before and after density modification.

<sup>‡</sup> R-factor =  $\Sigma(|F_{\text{obs}}| - k|F_{\text{calc}}|) / \Sigma |F_{\text{obs}}|$  and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.



Extended Data Table 2 | Oligonucleotides used in this study

Primer Name	5' Sequence 3'
NisB R784A F	GAT ATA CTG ATC CTA AAC CAC ATA TTG CAT TGC GTA TAA AAT GTT CAG ATT TAT TTT TA
NisB R784A R	AAT ATG TGG TTT AGG ATC AGT ATA TCT TAG GAA GAA TAG ATT TCC ACC CA
<i>L. lactis</i> <i>gltX</i> NdeI F	AAT ATA ATC ATA TGA ACA AAA AAA TCC GCG TC
<i>L. lactis</i> <i>gltX</i> XhoI R	AAT AAT ACT CGA GTT ATT TAT TTA AAG CGG CAT CCA AG
<i>E. coli</i> <i>gltX</i> NdeI F	AAT CAA TCA TAT GAA AAT CAA AAC TCG CTT CG
<i>E. coli</i> <i>gltX</i> XhoI R	TTA CTA CTC GAG TTA CTG CTG ATT TTC GCG TTC AG
<i>E. coli</i> tRNA <sup>Glu</sup> F	AAT TCC TGC AGT AAT ACG ACT CAC TAT AGT CCC CTT CGT CTA GAG GCC CAG GAC ACC
<i>E. coli</i> tRNA <sup>Glu</sup> R*	mUmGG CGT CCC CTA GGG GAT TCG AAC CCC TGT TAC CGC CGT GAA AGG GCG GTG TCC TGG

\*m stands for methylation at the 2' position of the ribose sugar

# Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation

Karl A. P. Payne<sup>1\*</sup>, Carolina P. Quezada<sup>1\*</sup>, Karl Fisher<sup>1</sup>, Mark S. Dunstan<sup>1</sup>, Fraser A. Collins<sup>1</sup>, Hanno Sjuts<sup>1†</sup>, Colin Levy<sup>1</sup>, Sam Hay<sup>1</sup>, Stephen E. J. Rigby<sup>1</sup> & David Leys<sup>1</sup>

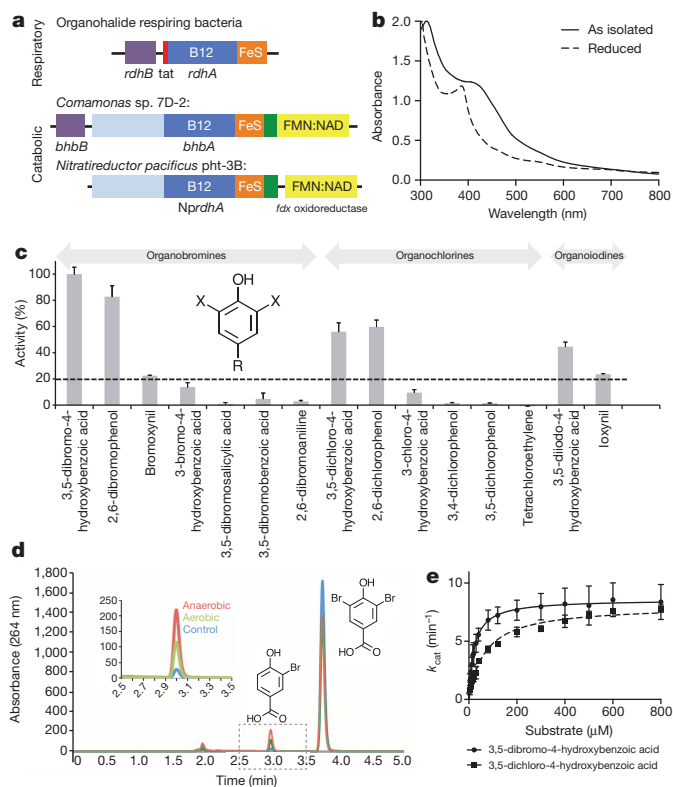
Organohalide chemistry underpins many industrial and agricultural processes, and a large proportion of environmental pollutants are organohalides<sup>1</sup>. Nevertheless, organohalide chemistry is not exclusively of anthropogenic origin, with natural abiotic and biological processes contributing to the global halide cycle<sup>2,3</sup>. Reductive dehalogenases are responsible for biological dehalogenation in organohalide respiring bacteria<sup>4,5</sup>, with substrates including polychlorinated biphenyls or dioxins<sup>6,7</sup>. Reductive dehalogenases form a distinct sub-family of cobalamin (B12)-dependent enzymes that are usually membrane associated and oxygen sensitive, hindering detailed studies<sup>8–12</sup>. Here we report the characterization of a soluble, oxygen-tolerant reductive dehalogenase and, by combining structure determination with EPR (electron paramagnetic resonance) spectroscopy and simulation, show that a direct interaction between the cobalamin cobalt and the substrate halogen underpins catalysis. In contrast to the carbon-cobalt bond chemistry catalysed by the other cobalamin-dependent subfamilies<sup>13</sup>, we propose that reductive dehalogenases achieve reduction of the organohalide substrate via halogen–cobalt bond formation. This presents a new model in both organohalide and cobalamin (bio)chemistry that will guide future exploitation of these enzymes in bioremediation or biocatalysis.

Cobalamin consists of a cobalt ion coordinated by a tetrapyrrole-derived macrocycle and readily forms a weak organometallic C–Co(III) bond<sup>14</sup>. This bond can be homolytically or heterolytically cleaved and the B12-dependent enzymes can be classified according to the type of reaction catalysed<sup>13</sup>: (1) the 5'-deoxyadenosyl cobalamin (AdoCbl) dependent isomerases that use AdoCbl as a reusable radical via homolytic cleavage of the Co–C bond; (2) the methylcobalamin-dependent methyltransferases that catalyse a heterolytic C–Co(III) cleavage, cycling between cob(I)alamin and methylcob(III)alamin redox states; (3) the reductive dehalogenases. Little is known about the role, if any, of C–Co chemistry in biological reductive dehalogenation.

After the screening of a range of reductive dehalogenase genes (*rdhA*) for soluble heterologous expression (including recently uncovered catabolic *rdhA* family members<sup>15</sup>), active protein could be obtained when expressing a reductive dehalogenase homologue from *Nitratireductor pacificus* pht-3B (*NprdhA*) in *Bacillus megaterium*. Heterologous expression in *Escherichia coli* did yield soluble NpRdhA, but the enzyme lacked cobalamin. In contrast to most *rdhA* homologues, *NprdhA* lacks both the twin-arginine signal and the associated transmembrane *rdhB* gene, suggesting that it is a soluble cytoplasmic enzyme (Fig. 1a). *NprdhA* is closely related in sequence to catabolic reductive dehalogenases<sup>15</sup>, and more distantly to the respiratory chlorophenol reductases (Extended Data Fig. 1).

NpRdhA could be purified under aerobic conditions without detrimental effect on cofactor content or activity, and the UV–visible spectrum of purified NpRdhA is similar to that previously reported for respiratory RdhA proteins<sup>8–12</sup>, with a distinct feature at 390 nm for the reduced sample that probably corresponds to the cob(I)alamin species

(Fig. 1b). EPR spectroscopy of aerobically purified NpRdhA in frozen solution reveals a five-coordinate 'base off' cob(II)alamin cofactor, while reduction of the sample leads to a  $2 \times [4\text{Fe}–4\text{S}]^{1+}$  signal (Extended Data Fig. 2). This is similar to the membrane-associated RdhA enzymes. Spin quantification leads to 0.82 Co(II) and 1.84 reduced 4Fe–4S clusters per NpRdhA, in agreement with Co(I) estimation (Fig. 1b) and Fe



**Figure 1 | Characterization of NpRdhA as an *ortho*-dibromophenol reductase.** **a**, Cartoon representation of the domain structure of the two reductive dehalogenase enzyme classes. **b**, UV–visible spectrum of 79 μM NpRdhA as purified under aerobic conditions and following reduction using deazaflavin and EDTA. The Co(I) concentration estimated using  $\epsilon = 26,000 \text{ M}^{-1} \text{ cm}^{-1}$  suggests 0.8 Co per NpRdhA. **c**, Relative activity with organohalide substrates using methyl viologen as electron donor. Results are shown as mean  $\pm$  s.e.m.;  $n = 2$ . Highest activities (above the dotted line) were obtained when substrates resembled the 2,6-dihalophenol structure shown as inset. **d**, HPLC product profiles obtained when using spinach ferredoxin:NADP<sup>+</sup> oxidoreductase and ferredoxin to couple NADPH oxidation to NpRdhA 3,5-dibromo-4-hydroxybenzoic acid reductase activity. **e**, Steady-state kinetic profile obtained with 3,5-dibromo-4-hydroxybenzoic acid or 3,5-dichloro-4-hydroxybenzoic acid and reduced methyl viologen as substrates. Data points are shown as mean  $\pm$  s.d.;  $n = 4$ .

<sup>1</sup>Manchester Institute for Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK. <sup>†</sup>Present address: Institute of Biochemistry, Goethe Universität Frankfurt, Max-von-Laue-Straße 9, 60438 Frankfurt, Germany.

\*These authors contributed equally to this work.

quantification using bathophenanthroline. The enzyme catalyses the reduction of *ortho*-halogenated phenolic compounds using either reduced methyl viologen or NADPH (via spinach ferredoxin and ferredoxin:NADP<sup>+</sup> oxidoreductase) as electron donors (Fig. 1c). The latter reaction can be performed under aerobic conditions (Fig. 1d). NpRdhA displays a strict requirement for *o*-halogenated phenolic substrates, with a clear preference for those containing a 2,6-dihalogenated phenolic moiety. Highest activity was observed with 3,5-dibromo-4-hydroxybenzoic acid—a breakdown product of the herbicide bromoxynil (Fig. 1e).

To provide mechanistic insights and a structural basis for the observed substrate specificity, we determined the 2.3 Å crystal structure of NpRdhA (Extended Data Table 1). The structure reveals a compact globular fold with a central core domain responsible for cofactor binding (residues 244–606) that corresponds in size to the respiratory RdhA enzymes (Fig. 2a). The cobalamin cofactor is bound in the base-off form by a flavin-reductase-like domain (residues 244–505), reminiscent of the human vitamin-B12-processing enzyme CblC, the closest structural homologue available<sup>16</sup> (Extended Data Fig. 3a). The amino-terminal

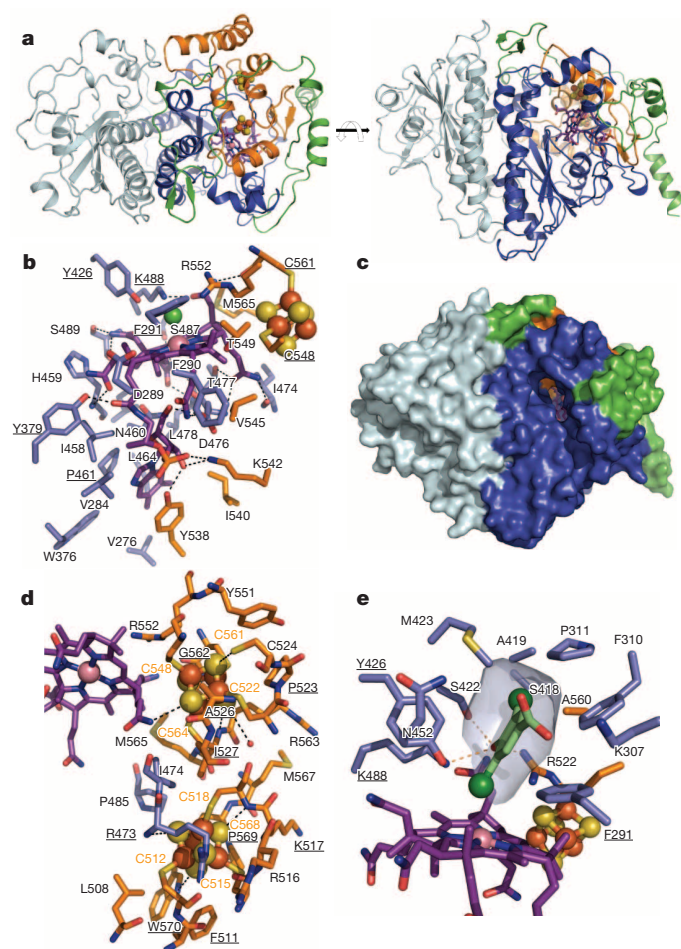
domain (residues 1–243) represents a duplication and divergence of the cobalamin-binding domain.

The cobalamin ligand is located at the extensive interface between the flavin-reductase domain and the carboxy-terminal 2×[4Fe–4S] binding region, and is bound by 11 direct hydrogen bonds between the corrin ring and the protein (Fig. 2b). The base-off conformation of 5,6-dimethylbenzimidazole is stabilized by four direct hydrogen bonds (in addition to a network of water-mediated polar interactions), an ionic interaction between the cobalamin phosphate and Lys 542, and the stacking of the 5,6-dimethylbenzimidazole moiety in between Pro 461 and Tyr 538. The C $\beta$  of Asp 476 is located directly below the Co ion (~5.5 Å), hindering binding of any solvent-derived ligands. In contrast, the opposite side of the corrin ring is located at the bottom of a letter-box-shaped cavity providing direct access to solvent (Fig. 2c).

The C-terminal region (506–694) consists of a series of  $\alpha$ -helices and extended loop regions that wrap around the functional cobalamin-binding domain. The 505–606 stretch contains the bacterial 2×[4Fe–4S] ferredoxin-like motif and binds both [4Fe–4S] clusters, which are held approximately 9.8 Å apart (closest Fe–Fe distance) and are both located near the surface of the protein (Fig. 2a, d). The first [4Fe–4S] cluster is linked by the first three cysteines of a typical bacterial CXXCXXCX XXCP ferredoxin motif and located approximately 12.4 Å from the cobalamin (nearest S–corrin ring distance). The second [4Fe–4S] cluster is located in direct van der Waals contact with the edge of B pyrrole from the corrin ring (Fig. 2d). In the latter case, the corresponding [4Fe–4S] binding motif contains a variable insertion between the first (Cys 548) and second cysteine (Cys 561) in the majority of reductive dehalogenases. Despite the fact that all cysteine ligands are conserved, this cluster is replaced in some reductive dehalogenases by a [3Fe–4S] cluster<sup>8–12</sup>. Of note, Cys 548 is in direct van der Waals contact with the cobalamin (Extended Data Fig. 3b); with few exceptions<sup>17</sup> Fe–S clusters are rarely in direct contact with porphyrin cofactors. This is suggestive of a link between FeS cluster maturation and cobalamin incorporation in reductive dehalogenases.

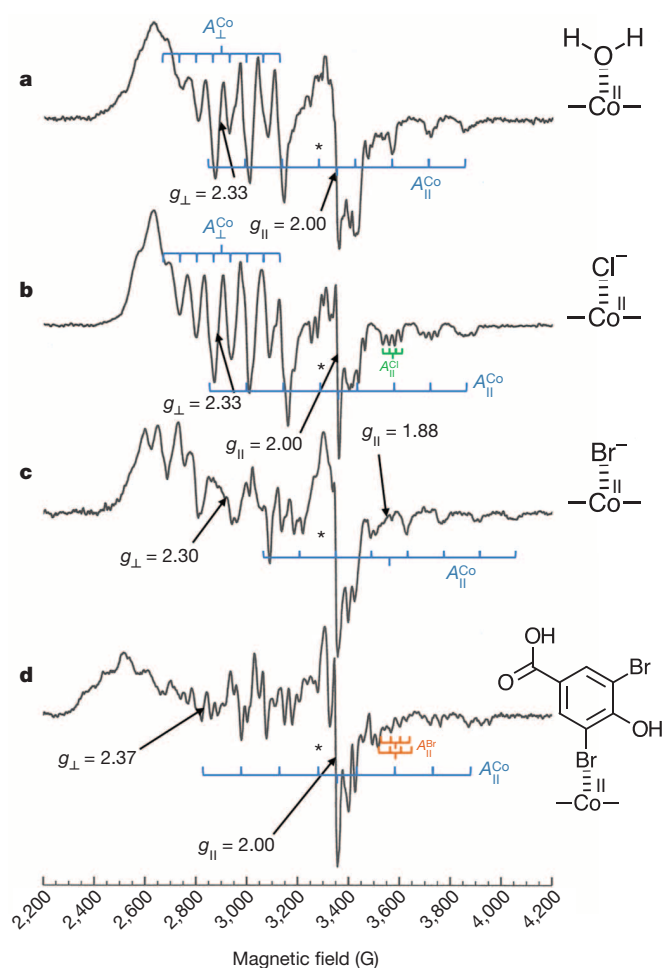
The reductive dehalogenase active site can clearly be identified as the solvent-exposed cavity above the cobalamin plane (Fig. 2b, c). While we have been unable to determine the structure of a substrate–enzyme complex, docking of the relatively rigid 3,5-dibromo-4-hydroxybenzoic acid substrate leads to a plausible model whereby one of the substrate bromide atoms is located directly above the Co ion (Fig. 2e), in close contact with the side chains of Tyr 426, Lys 488 and Arg 552. The substrate aromatic plane is orientated near perpendicular to the corrin ring, with the second bromide atom located in a hydrophobic cavity. The substrate hydroxyl group is in hydrogen-bonding contact with Ser 422, Lys 488 and Arg 552, suggesting that it is bound in the deprotonated state. Close inspection of the electron density corresponding to a fifth ligand to the Co ion reveals that it is best modelled by a chloride ion (Fig. 2b and Extended Data Fig. 3c). Furthermore, low-resolution NpRdhA structures of iodide-containing crystals display difference densities at locations that correspond to those of the bromide atoms in the modelled substrate–enzyme complex (Extended Data Fig. 3d).

EPR spectroscopic studies confirm the presence of a halogen–cobalamin interaction predicted by the substrate–enzyme model. Spectra of the type observed in Fig. 3a have been reported for various AdoCbl-dependent proteins<sup>18</sup> and for cobinamides<sup>19</sup> in solution, and typically arise from cob(II)alamin with water as the fifth axial ligand. Addition of NaCl to NpRdhA gave rise to the spectrum in Fig. 3b. While superficially resembling that of Fig. 3a, there are additional features in Fig. 3b arising from interaction of the unpaired electron with an additional nucleus having nuclear spin quantum number  $I = 3/2$ , resulting in a quartet of four lines. This indicates an unexpected interaction of the unpaired electron with a coordinated chloride ion ( $A_{\parallel}^{\text{Cl}}$ ) (where  $A$  is a component of the anisotropic hyperfine coupling constant). The inclusion of KBr in place of NaCl gave the spectrum in Fig. 3c. Bromide interacting with the cob(II)alamin should give rise to two overlapped quartets with larger hyperfine splitting than is observed for Cl. However, the spectrum



**Figure 2 | Crystal structure of NpRdhA in the resting state.** **a**, Overall structure of NpRdhA colour coded according to domain structure as in Fig. 1a and presented in two orientations. **b**, Detailed view of the cobalamin-binding pocket. Key residues are shown with carbons coloured as in panel **a**. Direct polar interactions between cobalamin and the enzyme are indicated by black dashed lines. Residues that are conserved in an alignment of functionally characterized RdhA enzymes are underlined. **c**, Solvent-accessible surface of NpRdhA colour coded as in panel **a**; the orientation is similar to that of the right-hand panel in **a**. **d**, Detailed view of the 2×[4Fe–4S] cluster binding region, representation as in panel **b**. **e**, Detailed view of the NpRdhA active site with a docked 3,5-dibromo-4-hydroxybenzoic acid substrate. Representation as in panel **b**; a transparent surface indicates the hydrophobic cavity that serves to bind the non-cobalt ligating bromine.





**Figure 3 | EPR spectroscopy of NpRdhA reveals a direct halogen–cobalamin interaction.** **a–d**, X-band (~9.39 GHz) continuous wave of 150  $\mu$ M NpRdhA in non-halide-containing buffer with  $g$  values marked (**a**); NpRdhA plus 200 mM NaCl with a ‘quartet’ arising from chloride ( $^{35}\text{Cl}$ ) superhyperfine coupling indicated and  $g$  values marked (**b**); NpRdhA plus 25 mM KBr with  $g$  values marked (**c**); NpRdhA plus 20 mM 3,5-dibromo-4-hydroxybenzoic acid with an example of the overlapped quartets arising from superhyperfine coupling to  $^{79}\text{Br}$  and  $^{81}\text{Br}$  indicated and  $g$  values marked (**d**). Cobalt hyperfine coupling is indicated in blue (showing the eight expected for an  $I = 7/2$   $^{59}\text{Co}$ ); halogen superhyperfine coupling (where present) is also indicated. For spectra **c** and **d** that exhibit overlapped  $A_{\perp}^{\text{Co}}$  and  $A_{\parallel}^{\text{Co}}$  only  $A_{\parallel}^{\text{Co}}$  is indicated. For spectrum parameters, see Extended Data Table 3. Asterisk indicates a superposition of a  $[3\text{Fe-4S}]^+$  spectrum, the spectrum of a cob(III)alamin-superoxide species and an unassigned organic radical signal which together comprise less than 5% of the cob(II)alamin spectrum (see Extended Data Fig. 4).

we observe displays a complex mixture of Co and Br (super)hyperfine couplings ( $A_{\perp}^{\text{Co}}$  and  $A_{\parallel}^{\text{Br}}$ ) around  $g_{\perp}$  and an unusually low  $g_{\parallel}$  (where  $g$  is a component of the anisotropic  $g$  tensor) with no detectable Br coupling. The same phenomenon has been reported for the iodo-cob(II)alamin complex in methanol<sup>18</sup> and is a clear indication of a Br–cob(II)alamin interaction. Addition of the substrate 3,5-dibromo-4-hydroxybenzoic acid produces the spectrum in Fig. 3d, with overlapping contributions from Co and Br (super)hyperfine coupling around  $g_{\perp}$  ( $A_{\perp}^{\text{Co}}$  and  $A_{\parallel}^{\text{Br}}$ ) and also a complex, overlapped pattern of Br superhyperfine coupling in the parallel features ( $A_{\parallel}^{\text{Br}}$ ) that indicates unprecedented interaction of the unpaired electron on cob(II)alamin with a Br atom of the substrate. Addition of weaker binding substrates such as 3,5-dichloro-4-hydroxybenzoic acid gave rise to multiple species reflecting mobility within the active site (Extended Data Fig. 4).

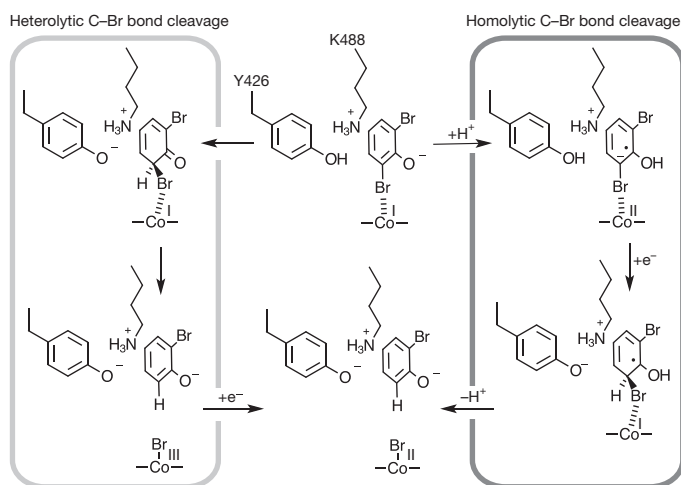
To probe the effects of halogen–cobalt coordination further, we carried out density functional theory (DFT) calculations on an active site

model (Extended Data Fig. 5 and Supplementary Table 1). These calculations suggest that substrate binding is likely to be stabilized by weak Co(II)–Br ligation and through hydrogen bonding of Lys 488 and Arg 552 to the substrate hydroxyl group. While the cobalamin–Br ligation is expected to be lost upon reduction to the Co(I) state, this is accompanied by an elongation of the substrate C–Br bond and a contraction of the Co–Br distance from 2.84 Å to 2.72 Å (Extended Data Table 2).

Previous suggestions for the mechanism of reductive dehalogenase catalysis can be roughly divided into those invoking a transient organo-cobalt adduct and those favouring a radical mechanism<sup>13</sup>. The NpRdhA structure is incompatible with formation of an organocobalt adduct (Extended Data Fig. 3e). The proposed radical mechanisms place fewer restraints on the relative position of Co and substrate, and therefore appear compatible with our structural data, although no direct evidence for a substrate radical has been obtained for RdhA enzymes. However, our structural and spectroscopic data suggest an alternative route, whereby the cobalamin is directly involved in halogen abstraction through formation of a halogen–cobalt bond.

Formation of a halogen–cob(II)alamin bond (as detected by EPR) drives catalysis, with the substrate bromide atom replacing the fifth ligand before reductive dehalogenation. In this configuration, the conserved Tyr 426/Lys 488 dyad is ideally placed to act as a hydrogen donor to the halogenated carbon (Fig. 4). We propose that the carbon–halogen bond could break either heterolytically, with transient formation of a halogen–cob(III)alamin following nucleophilic attack by Co(I), or homolytically, following formation of an aryl radical through electron transfer from either the Co(I) or the adjacent  $[4\text{Fe-4S}]$  cluster. We show that mutation of Y426F or K488Q does not appear to affect corrinoid or  $[4\text{Fe-4S}]$  binding but abolishes enzyme activity, implicating both residues in the RdhA catalytic mechanism (Extended Data Fig. 6). While the R552L NpRdhA variant also lacks activity, this variant has markedly altered properties.

Our proposed NpRdhA mechanism is fundamentally different from other B12-containing enzymes<sup>13,14</sup> and from the hydrolytic dehalogenases, which use a direct  $\text{S}_{\text{N}}2$ -type attack on the substrate carbon<sup>20</sup>. In contrast, we propose that the reductive dehalogenase uses the cobalamin cofactor to attack the substrate halogen atom itself, leading to C–halogen bond breakage concomitant with protonation of the leaving group. Distinct variations on this theme could occur within the RdhA family to account for the varied substrate specificity (aliphatic versus aromatic organohalides) as well as the dihalo-elimination catalysed by



**Figure 4 | NpRdhA proposed mechanism.** Our data suggest that the NpRdhA mechanism involves formation of a bromide–cob(II)alamin complex either via heterolytic cleavage of the carbon–bromide substrate bond (left box) or homolytic cleavage of the carbon–bromide substrate bond (right box). For clarity, only the 2,6-dibromophenol moiety of the substrate is shown.



some enzymes<sup>21</sup> (Extended Data Fig. 7). The formation of a cobalamin-halide complex via oxidative addition to the Co ion is a new paradigm in both organohalide and cobalamin biochemistry and suggests that the B12 biochemical repertoire is probably more varied than previously anticipated. In fact, recent studies hint at a distant enzyme relative being involved in tRNA modification<sup>22</sup>. Our findings should serve to guide future exploitation of the reductive dehalogenases.

**Note added in proof:** The crystal structure of a respiratory reductive dehalogenase, PceA from *Sulfurospirillum multivorans*, shows remarkable similarities to NpRdhA. A iodide-cob(II)alamin complex was also observed, while ligand-enzyme complexes revealed several conformations for the substrate or product within the active site (ref. 23).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 July; accepted 30 September 2014.

Published online 19 October 2014.

1. Stringer, R. & Johnston, P. *Chlorine and the Environment. An Overview of the Chlorine Industry* (Kluwer Academic, 2001).
2. Öberg, G. The natural chlorine cycle—fitting the scattered pieces. *Appl. Microbiol. Biotechnol.* **58**, 565–581 (2002).
3. Gribble, G. W. Occurrence of halogenated alkaloids. *Alkaloids Chem. Biol.* **71**, 1–165 (2012).
4. Smidt, H. & de Vos, W. M. Anaerobic microbial dehalogenation. *Annu. Rev. Microbiol.* **58**, 43–73 (2004).
5. Leys, D., Adrian, L. & Smidt, H. Organohalide respiration: microbes breathing chlorinated molecules. *Phil. Trans. R. Soc. Lond. B* **368**, 20120316 (2013).
6. Adrian, L., Szwed, U., Wecke, J. & Görisch, H. Bacterial dehalorespiration with chlorinated benzenes. *Nature* **408**, 580–583 (2000).
7. Bunge, M. *et al.* Reductive dehalogenation of chlorinated dioxins by an anaerobic bacterium. *Nature* **421**, 357–360 (2003).
8. Wohlfarth, G. & Diekert, G. Anaerobic dehalogenases. *Curr. Opin. Biotechnol.* **8**, 290–295 (1997).
9. van de Pas, B. A. *et al.* Purification and molecular characterization of ortho-chlorophenol reductive dehalogenase, a key enzyme of halo-respiration in *Desulfotobacterium dehalogenans*. *J. Biol. Chem.* **274**, 20287–20292 (1999).
10. Krasotkina, J., Walters, T., Maruya, K. A. & Ragsdale, S. W. Characterization of the B12- and iron-sulfur-containing reductive dehalogenase from *Desulfotobacterium chlororespirans*. *J. Biol. Chem.* **276**, 40991–40997 (2001).
11. Neumann, A. *et al.* Tetrachloroethene reductive dehalogenase of *Dehalospiroillum multivorans*: substrate specificity of the native enzyme and its corrinoid cofactor. *Arch. Microbiol.* **177**, 420–426 (2002).
12. Maillard, J. *et al.* Characterization of the corrinoid iron-sulfur protein tetrachloroethene reductive dehalogenase of *Dehalobacter restrictus*. *Appl. Environ. Microbiol.* **69**, 4628–4638 (2003).
13. Banerjee, R. & Ragsdale, S. W. The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annu. Rev. Biochem.* **72**, 209–247 (2003).
14. Brown, K. L. Chemistry and enzymology of vitamin B12. *Chem. Rev.* **105**, 2075–2150 (2005).
15. Chen, K. *et al.* Molecular characterization of the enzymes involved in the degradation of a brominated aromatic herbicide. *Mol. Microbiol.* **89**, 1121–1139 (2013).
16. Koutmos, M., Gherasim, C., Smith, J. L. & Banerjee, R. Structural basis of multifunctionality in a vitamin B12-processing enzyme. *J. Biol. Chem.* **286**, 29780–29787 (2011).
17. Stroupe, M. E. & Getzoff, E. D. In *Tetrapyrroles: Their Birth, Life, and Death* (eds Warren, M. J. & Smith, A.) 373–387 (Landes Bioscience, 2005).
18. Van Doorslaer, S. *et al.* Axial solvent coordination in “base-off” cob(II)alamin and related co(II)-corrinates revealed by 2D-EPR. *J. Am. Chem. Soc.* **125**, 5915–5927 (2003).
19. Bayston, J. H., Looney, F. H., Pilbrow, J. R. & Winfield, M. E. Electron paramagnetic resonance studies of cob(II)alamin and cob(II)inamides. *Biochemistry* **9**, 2164–2172 (1970).
20. de Jong, R. M. & Dijkstra, B. W. Structure and mechanism of bacterial dehalogenases: different ways to cleave a carbon-halogen bond. *Curr. Opin. Struct. Biol.* **13**, 722–730 (2003).
21. Grostern, A. & Edwards, E. A. Characterization of a *Dehalobacter* coculture that dechlorinates 1,2-dichloroethane to ethene and identification of the putative reductive dehalogenase gene. *Appl. Environ. Microbiol.* **75**, 2684–2693 (2009).
22. Miles, Z. D., McCarty, R. M., Molnar, G. & Bandarian, V. Discovery of epoxycyclohexanone (oQ) reductase reveals parallels between halo-respiration and tRNA modification. *Proc. Natl Acad. Sci. USA* **108**, 7368–7372 (2011).
23. Bommer, M. *et al.* Structural basis for organohalide respiration. *Science* <http://dx.doi.org/10.1126/science.1258118> (2 October 2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by an ERC grant (DEHALORES206080) to D.L. S.H. is a BBSRC David Phillips research fellow. C.P.Q. is supported by CONICYT Chile. We thank Diamond Light Source for access to beamline I04 (proposal number MX8997) that contributed to the results presented here; R. Biedendieck (Technische Universität Braunschweig, Germany), S. Moore and M. Warren (University of Kent, UK) for providing plasmids and assistance with *B. megaterium* molecular biology; the assistance given by IT Services and the use of the Computational Shared Facility at The University of Manchester; and A. Munro (University of Manchester) for critical comments and discussion.

**Author Contributions** K.A.P.P., M.S.D., K.F. and H.S. were involved in *rdhA* heterologous production screening. K.A.P.P. obtained heterologous expression of active NpRdhA in *B. megaterium*. C.P.Q. crystallized and solved the NpRdhA structure with assistance from D.L. and C.L.; C.P.Q., K.A.P.P., M.S.D. and K.F. performed kinetic studies; K.A.P.P. and F.A.C. prepared and analysed NpRdhA variants. K.F. and S.E.J.R. performed and analysed EPR experiments. S.H. performed docking and DFT calculations. All authors discussed the results and participated in writing the manuscript. D.L. initiated and directed this research.

**Author Information** Coordinates and structure factors have been deposited in the Protein Data Bank under accession number 4RAS. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.L. ([david.leys@manchester.ac.uk](mailto:david.leys@manchester.ac.uk)).

## METHODS

### Cloning of *N. pacificus* NprdhA for *B. megaterium* heterologous expression.

A BLASTp search of the *Nitratireductor pacificus* pht-3B genome<sup>24</sup> using the protein sequence of *Comamonas* sp. 7D-2 BbbA (YP\_007878394) revealed three putative reductive dehalogenase homologues sharing 72% (WP\_008593084.1), 51% (WP\_008595008.1) and 42% (WP\_008597722.1) identity with BbbA. The *N. pacificus* WP\_008597722.1 gene (NprdhA) was shorter than the other homologues, lacking the putative C-terminal reductase domain. The NprdhA gene was codon optimized to remove codons that were rare in both *E. coli* and *B. megaterium* and synthesized (Genscript). *N. pacificus* NprdhA was PCR amplified using primers NPRdhA3pPT7F and NPRdhA3pPT7R (the latter containing the sequence encoding for a C-terminal His tag) using Phusion polymerase (NEB). The PCR product was cloned into the BsrGI and BamHI sites of pP<sub>T7</sub> plasmid<sup>25</sup> using Infusion HD (Clontech) and transformed into *E. coli* NEB5 $\alpha$ . Once the sequence of the insert was confirmed the purified plasmid was transformed into *B. megaterium* MS941 containing the pT7-RNAP plasmid that permits xylose-inducible expression of T7 polymerase<sup>24</sup>, using the modified minimal medium protoplast transformation protocol<sup>26</sup>.

**Mutagenesis.** Mutagenesis primers were designed using the QuikChange Primer Design Program (<http://www.genomics.agilent.com/primerDesignProgram.jsp>). PCR was performed using Phusion polymerase (NEB). Template was removed by DpnI (NEB) digest and the PCR product transformed into *E. coli* NEB5 $\alpha$ . Once the presence of the desired mutation was confirmed by DNA sequencing, the plasmid was transformed into *B. megaterium* as above.

### Heterologous expression of NprdhA wild-type and mutants in *B. megaterium*.

*B. megaterium* transformants were grown in 1 l LB supplemented with 10  $\mu\text{g ml}^{-1}$  tetracycline and 4.5  $\mu\text{g ml}^{-1}$  chloramphenicol at 37 °C, 180 r.p.m. until the culture reached an OD<sub>578</sub> of ~0.4. Cultures were supplemented with 50  $\mu\text{M}$  ammonium iron(II) sulphate, 1  $\mu\text{M}$  B12 and induced with 0.1% xylose. Cultures were grown overnight at 17 °C, 180 r.p.m. and then harvested by centrifugation (4 °C, 7,000g for 10 min).

**Purification of NpRdhA wild-type and variant enzymes.** Cell pellets were resuspended in buffer A (200 mM NaCl, 50 mM Tris pH 7.5) supplemented with DNase, RNase, lysozyme (Sigma) and a Complete EDTA-free protease inhibitor cocktail (Roche). Cells were lysed using a French press at 1,500 psi and the lysate clarified by centrifugation at 125,000g for 90 min. The supernatant was applied to a Ni-NTA agarose column (Qiagen). The column washed successively with three column volumes of buffer A supplemented with 10 mM and 40 mM imidazole and protein eluted in 1 ml fractions with buffer A supplemented with 250 mM imidazole. Samples were subjected to SDS-PAGE analysis and fractions found to contain the purified protein were pooled. Imidazole was removed using a 10-DG desalting column (Bio-Rad) equilibrated with buffer A. Protein was concentrated as required using a Vivaspin centrifugal device (Sartorius).

**UV-visible spectroscopy/protein quantification.** UV-visible absorbance spectra were recorded with a Cary UV-Vis spectrophotometer. To obtain the reduced NpRdhA spectrum, 79  $\mu\text{M}$  enzyme was incubated anaerobically with 25  $\mu\text{M}$  5-deazariboflavin and 2 mM EDTA under a blue LED lamp for 1 h. Protein concentration was estimated using  $\epsilon_{280} = 77,810 \text{ M}^{-1} \text{ cm}^{-1}$  (calculated from the primary amino acid sequence using the ProtParam program on the ExPASy proteomics server).

**Metal estimations.** The iron content of RdhA was confirmed colourimetrically with bathophenanthroline after acid denaturation. Protein was mixed with an equal volume of 2 M HCl and heated at 80 °C for 10 min before the removal of the precipitate by centrifugation. A suitable volume of the supernatant was taken for analysis and made up to a final volume of 0.75 ml with water. Samples were mixed after addition of 0.2 ml saturated ammonium acetate. Sodium metabisulphite (0.1 M, 0.05 ml) and bathophenanthroline (0.1 M, 0.01 ml) were added and samples incubated at room temperature for 10 min before reading the absorbance at 535 nm. Concentrations were determined from an iron standard curve measured over the range 0–50 nmol.

**Methyl viologen spectrophotometric activity assay.** Purified NpRdhA was buffer exchanged into 50 mM Tris, pH 7.5, 200 mM NaCl under anaerobic conditions in a glove box (Belle Technology, UK). A 30 mM methyl viologen solution reduced by titration with sodium dithionite was prepared for the assay. The substrate-dependent oxidation of dithionite-reduced methyl viologen was measured at 578 nm ( $\epsilon = 978 \text{ M}^{-1} \text{ cm}^{-1}$ ). Screening of different organohalide compounds was performed by adding 100  $\mu\text{M}$  of the substrate to the reaction mixture containing 2  $\mu\text{M}$  enzyme in 50 mM Tris, pH 7.5, 200 mM NaCl and 300  $\mu\text{M}$  dithionite-reduced methyl viologen. Oxidation of methyl viologen was measured for a period of 10 min and fit to a linear rate where appropriate. Full steady-state characterization was done for 3,5-dibromo-4-hydroxybenzoic acid or 3,5-dichloro-4-hydroxybenzoic acid respectively, and a plot of the observed rate constant versus substrate concentration was fitted to a Michaelis–Menten steady-state model.

**NpRdhA spinach ferredoxin and ferredoxin:NADP<sup>+</sup> oxidoreductase assay.** Reactions using spinach ferredoxin and ferredoxin:NADP<sup>+</sup> oxidoreductase were carried out either in an anaerobic glove box (Belle Technology, UK) or under aerobic

conditions. For anaerobic assays purified NpRdhA was buffer exchanged into 50 mM Tris, pH 7.5, 200 mM NaCl under strict anaerobic conditions. Reactions contained 100 nM NpRdhA, 200  $\mu\text{M}$  NADPH, spinach ferredoxin and ferredoxin:NADP<sup>+</sup> oxidoreductase in an enzyme ratio of 1:10:2 (NpRdhA:ferredoxin:ferredoxin:NADP<sup>+</sup> oxidoreductase). All assays were initiated by the addition of 2 mM 3,5-dibromo-4-hydroxybenzoic acid. After incubation for 1 h, reactions were stopped by rapid freezing in liquid nitrogen and analysed by HPLC.

**HPLC assays.** Sample analysis was performed using an Agilent 1100 Series HPLC equipped with a UV detector. The stationary phase was a Kinetex 5  $\mu\text{m}$  C18 100A column, 250  $\times$  4.6 mm. The mobile phase was acetonitrile/water (50/50) with 0.1% TFA at a flow rate of 1 ml min<sup>-1</sup>, and detection was performed at a wavelength of 264 nm.

**Crystallization.** Purified NpRdhA protein in 50 mM Tris, pH 7.5, 200 mM NaCl was concentrated in a Vivaspin 30-kDa cutoff spin concentrator to a final concentration of 11 mg ml<sup>-1</sup>. Initial screening mixing 0.25  $\mu\text{l}$  protein with 0.25  $\mu\text{l}$  mother liquor led to small crystals in a variety of conditions when incubated at 4 °C. Crystals obtained with 0.2 M sodium iodide, 0.1 M Bis Tris propane, 20% w/v PEG3350 were used to prepare a seed stock, and a second screening was carried out by mixing 0.25  $\mu\text{l}$  protein plus 0.05  $\mu\text{l}$  seed stock with 0.3  $\mu\text{l}$  screen buffer. Crystals were obtained with 0.2 M potassium thiocyanate, 0.1 M sodium cacodylate pH 6.5, 8% w/v PEG 20K plus 8% v/v PEG550MME when incubating at 4 °C.

**Diffraction data collection and structure elucidation.** Crystals were flash-cooled in liquid nitrogen by supplementing the mother liquor with 10% PEG 200. Data were collected at Diamond beamlines at 100K and subsequently handled using the CCP4 suite<sup>27</sup>. All data were reduced and scaled using XDS<sup>28</sup> and initial phases obtained by merging various native data sets (all collected at 0.9795 Å wavelength) using BLEND<sup>27</sup>. This generated a highly redundant low-resolution data set with sufficient anomalous signal from the native Fe and Co ions to allow substructure determination using Shelx<sup>29</sup> and initial phasing using MLPHARE<sup>27</sup>. Interpretable maps were obtained following density modification and threefold NCS averaging combined with phase extension to 2.8 Å using DM<sup>27</sup>. An initial model was automatically generated using ArpWarp<sup>30</sup>, and iteratively rebuilt and refined against a single high-resolution data set using Coot and REFMAC5 (ref. 27). The final model was refined using data extending to 2.3 Å and contains three molecules in the asymmetric unit. While the identity of the fifth ligand to the Co ion could not be unambiguously established from the X-ray data alone, modelling this as a chloride ion resulted in the most plausible model and is in agreement with our EPR-spectroscopic data (when refining this as a water ligand, individual B factors for the water ligand were lower than that of the corresponding cobalt ion for all three monomers in the AU). A low-resolution data set (to 3.5 Å) of an NpRdhA crystal obtained in the presence of 200 mM sodium iodide was used to determine position of iodide sites using difference Fourier methods. For final data and refinement statistics see Extended Data Table 1. Ramachandran statistics of the model show 96.93% in the favoured and 2.54% in the allowed regions.

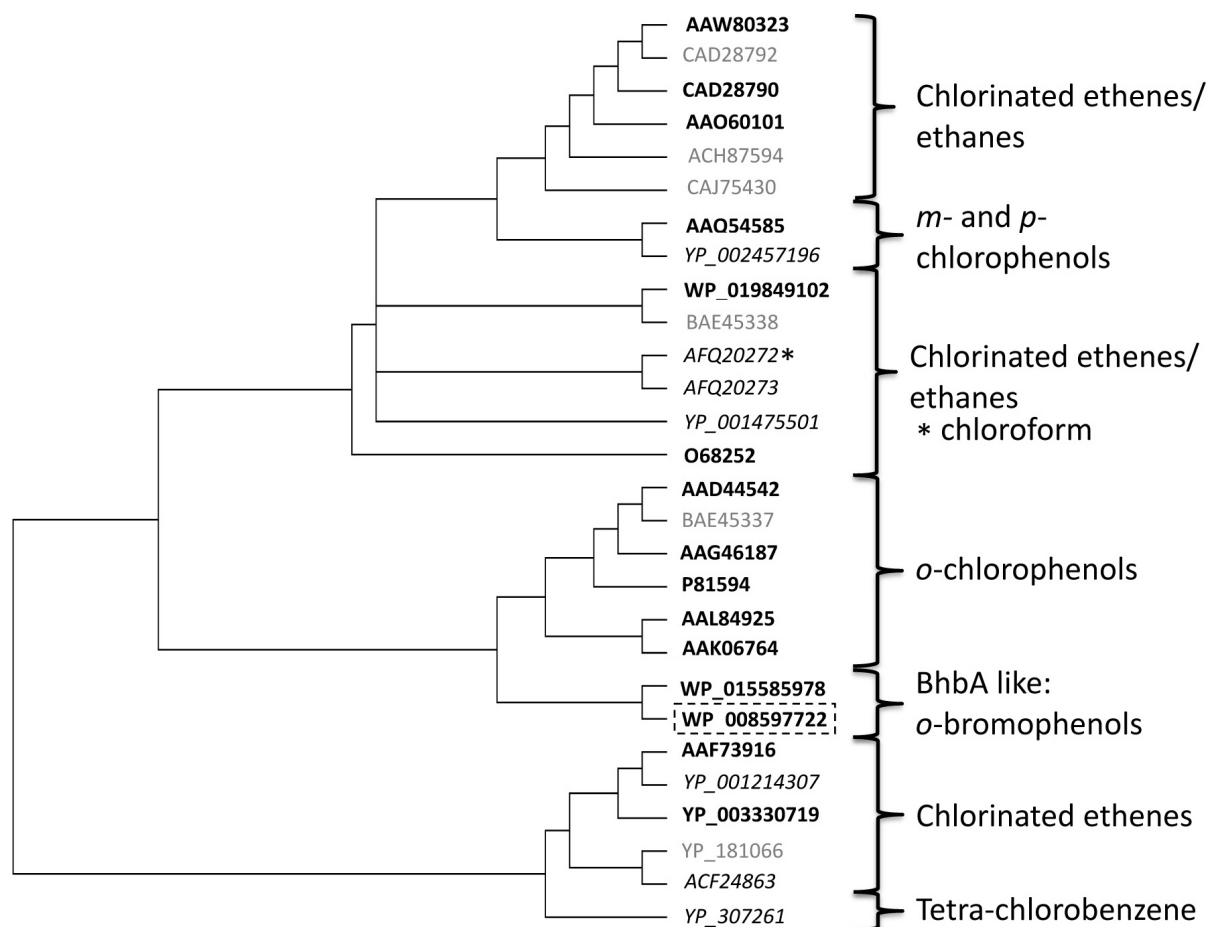
**EPR spectroscopy.** Experimental parameters: microwave power 0.5 mW, modulation frequency 100 kHz, modulation amplitude 5 G, temperature 30 K. Spectra were obtained using a Bruker ELEXSYS E500 spectrometer, Super high Q resonator (ER 4122SHQ), Oxford Instruments ESR900 cryostat and ITC503 temperature controller. Cob(II)alamin contains a low spin,  $S = \frac{1}{2}$  (that is, one unpaired electron) ion with electronic configuration [Ar]3d. The unpaired electron is in the  $d_{z^2}$  orbital and this leads to axial spectra that can be described by two  $g$  values and two cobalt hyperfine values ( $A^{\text{Co}}$ ), each designated parallel ( $\parallel$ ) and perpendicular ( $\perp$ ). The latter describe the (hyperfine) interaction of the single unpaired electron with the nuclear magnetic moment of the cobalt ion. The cobalt nucleus has nuclear spin  $I = 7/2$  and thus splits the spectrum into eight lines. In five coordinate base-off complexes as described here, the hyperfine and  $g$  tensors are not collinear in  $x$  and  $y$  and this together with dislocation strain leads to a lack of resolution in the  $A_{\perp}$  lines. The additional features in Fig. 3b arise from interaction of the unpaired electron with an additional nucleus having  $I = 3/2$ , giving rise to a quartet of four lines. This indicates an unprecedented interaction of the unpaired electron with a coordinated chloride ion; chlorine has  $I = 3/2$  and over 75% of natural chlorine is the <sup>35</sup>Cl isotope, leading to a chlorine superhyperfine interaction ( $A^{\text{Cl}}$ ). Bromine exists as two isotopes, <sup>79</sup>Br and <sup>81</sup>Br, in essentially a 1:1 ratio, each having  $I = 3/2$  but differing in magnetic moment by ~7.5% and with even the smaller of the two magnetic moments being ~2.5 times greater than that of Cl. Thus, bromide interacting with the cob(II)alamin should give rise to two overlapped quartets with larger hyperfine splitting than is observed for Cl. The concentrations of EPR-detectable Co(II) and reduced 4Fe–4S clusters were estimated by double integration of the X-band continuous wave EPR signals of these species recorded at non-saturating microwave powers and comparison of these double integrals with that of a Cu(II) standard of known concentration recorded under the same conditions. Error estimated at  $\pm 5\%$ .

**Docking.** Computational docking of the substrate 3,5-dibromo-4-hydroxybenzoic acid to NpRdhA was performed using AutoDock vina<sup>31</sup>. A doubly deprotonated

substrate model was optimized using the UFF force field with Gaussian 09 (ref. 32). AutoDock Tools 1.5.6 (ref. 33) was used to assign hydrogens and Gasteiger charges to NpRdhA. The substrate carboxy moiety was made rotatable, and the four residues Phe 291, Tyr 426, Lys 488 and Arg 552 were made flexible. A docking grid with dimensions of  $24 \times 36 \times 18 \text{ \AA}$  was used to include the entire active site/substrate binding cavity. The final docking conformation was chosen based on the conformation with the shortest substrate heavy atom to cobalamin Co distance. This conformation was  $1.4 \text{ kcal mol}^{-1}$  higher in energy than the highest scoring result.

**Density functional theory (DFT) calculations.** A DFT active site model based on the crystal structure of NpRdhA with docked substrate was optimized using the (U)BP86/6-31G(d,p) level of theory implemented in Gaussian 09 (ref. 32). The BP86 functional has been shown to be the most appropriate for the prediction of both structural and electronic properties of B12 cofactors (ref. 34 and references therein). The model consists of a 'trimmed' tetracoordinate cobalamin, the truncated side chains of Tyr 426, Lys 488 and Arg 552, and the substrate analogue 2,6-dibromo-4-methylphenol, which was deprotonated (Extended Data Fig. 5). Four atoms, including the cobalamin Co, were fixed during optimization. The Co(I), Co(II) and Co(III) oxidation states were examined and bromide-ligated cobalamin models were also examined for comparison of the Co–Br bond lengths.

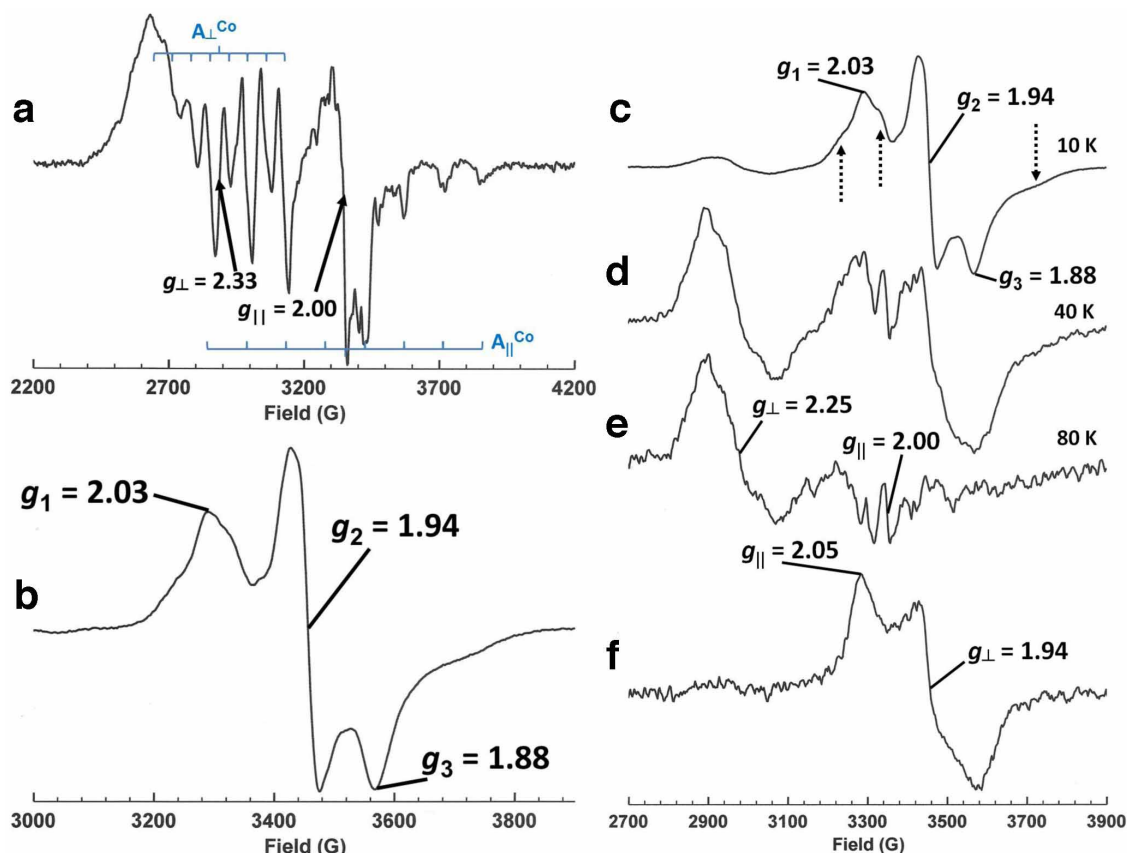
24. Lai, Q., Li, G. & Shao, Z. Genome sequence of *Nitratireductor pacificus* type strain pht-3B. *J. Bacteriol.* **194**, 6958 (2012).
25. Gamer, M. *et al.* A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.* **82**, 1195–1203 (2009).
26. Moore, S. J. *et al.* Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12). *Proc. Natl Acad. Sci. USA* **110**, 14906–14911 (2013).
27. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
28. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
29. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D* **66**, 479–485 (2010).
30. Langer, G. *et al.* Automated molecular model building for X-ray crystallography using ARP/Warp version 7. *Nature Protocols* **3**, 1171–1179 (2008).
31. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
32. Frisch, M. J. *et al.* Gaussian 09 (Gaussian, revision B.01, 2010).
33. Morris, G. M. *et al.* Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
34. Liu, H., Kornobis, K., Lodowski, P., Jaworska, M. & Kozlowski, P. M. TD-DFT insight into photodissociation of the Co–C bond in coenzyme B12. *Front. Chem.* **1**, 1–12 (2013).



**Extended Data Figure 1 | Phylogenetic tree of functionally characterized reductive dehalogenases (RdhAs).** Those in bold have been purified and characterized *in vitro*. Those in italics have been identified from crude lysate assays via native PAGE or knock out/knock in. Those in plain text have implied substrate range based on transcriptional regulation. The sequence accession codes are as follows: AAD44542, *Desulfitobacterium dehalogenans* ATCC 51507 CprA; P81594, *Desulfitobacterium hafniense* DCB-2 CprA1; AAQ54585, *Desulfitobacterium hafniense* PCP-1 CprA5; AAL84925, *Desulfitobacterium chlororespirans* CprA; WP\_019849102, *Desulfitobacterium* sp. PCE-1; AAK06764, *Desulfitobacterium hafniense* PCP-1 CprA3; AAO60101, *Desulfitobacterium hafniense* PCE-S PceA; O68252, *Dehalospirillum multivorans* PceA; AAF73916, *Dehalococcoides mccartyi* TceA; YP\_181066, *Dehalococcoides mccartyi* 195 PceA; AAW80323,

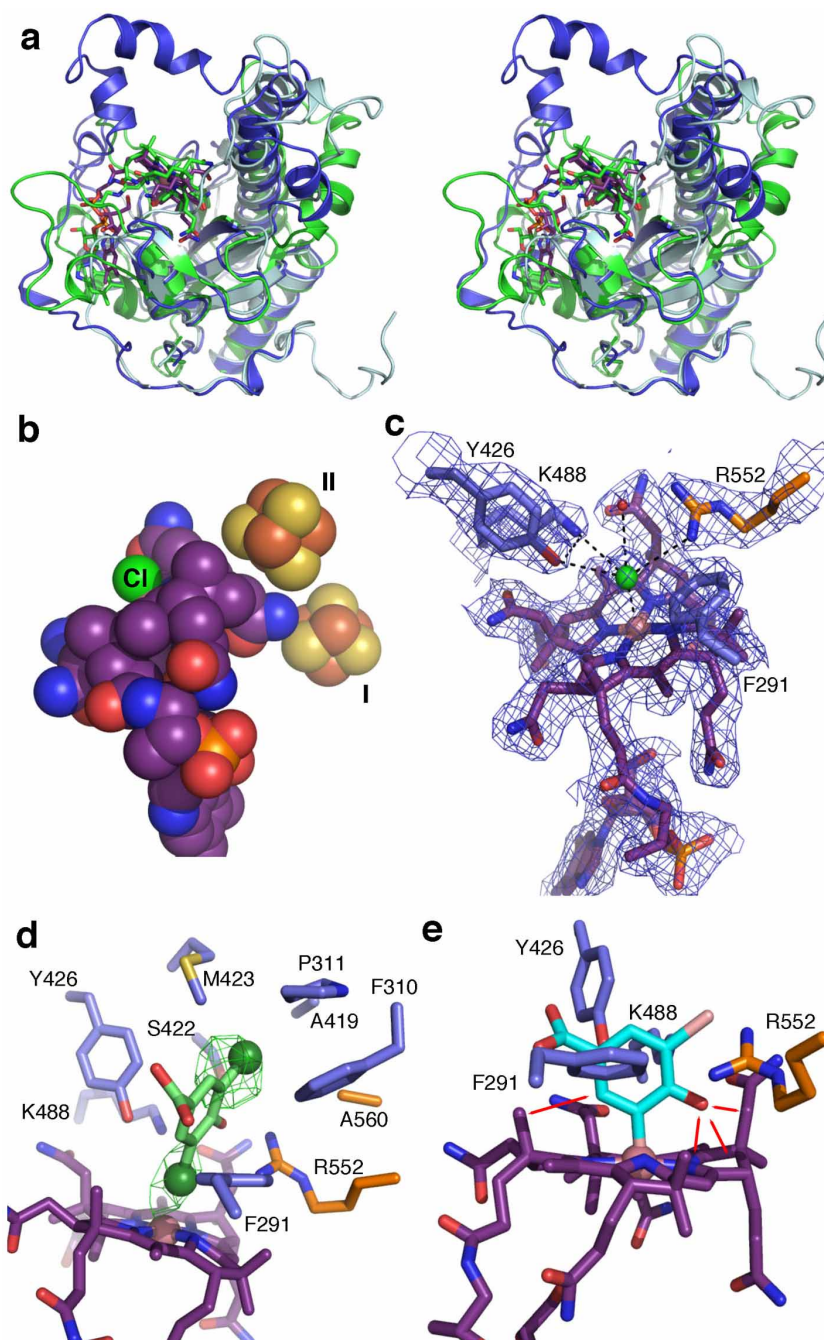
*Desulfitobacterium hafniense* Y51 PceA; CAD28790, *Dehalobacter restrictus* PceA; CAD28792, *Desulfitobacterium hafniense* TCE1 PceA; YP\_001214307, *Dehalococcoides mccartyi* BAV1 BvcA; YP\_003330719, *Dehalococcoides mccartyi* VS VcrA; ACF24863, *Dehalococcoides* sp. MB MbrA; YP\_307261, *Dehalococcoides mccartyi* CBDB1 CbrA; BAE45338, *Desulfitobacterium* sp. KBC1 PrdA; BAE45337, *Desulfitobacterium* sp. KBC1 CprA; ACH87594, *Dehalobacter* sp. WL DcaA; CAJ75430, *Desulfitobacterium dichloroeliminans* LMG P-21439 DcaA; AFQ20272, *Dehalobacter* sp. CF CfrA; AFQ20273, *Dehalobacter* sp. DCA DcrA; AAG46187, *Desulfitobacterium* sp. PCE1 CprA; YP\_002457196, *Desulfitobacterium hafniense* DCB-2 RdhA3; YP\_001475501, *Shewanella sediminis* HAW-EB3 PceA; WP\_015585978, *Comamonas* sp. 7D-2 BhbA; WP\_008597722, *Nitratireductor pacificus* NpRdhA.





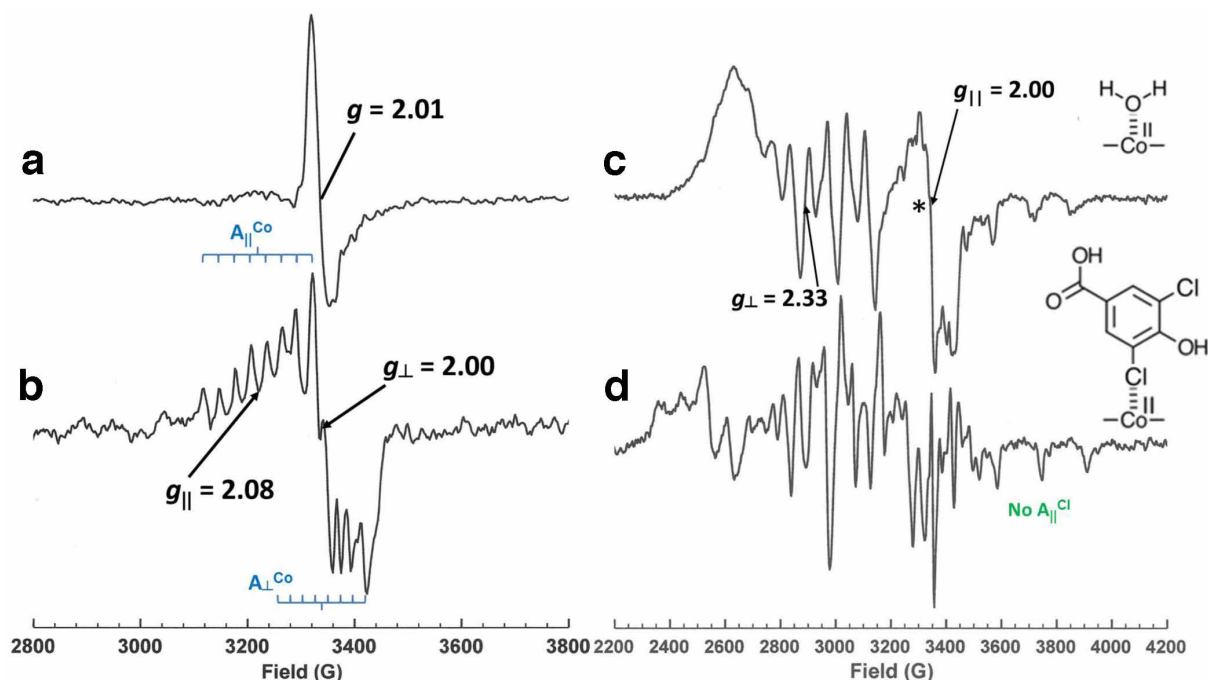
**Extended Data Figure 2 | EPR spectroscopic analysis of reduced NpRdhA (79  $\mu\text{M}$ ).** **a, b**, X-band EPR spectra of NpRdhA as isolated, 150  $\mu\text{M}$  (**a**), and reduced using EDTA, deazaflavin and light, 79  $\mu\text{M}$  (**b**). **c**, X-band ( $\sim 9.39$  GHz) continuous wave EPR spectrum of reduced NpRdhA recorded at 10 K showing a  $2 \times [4\text{Fe-4S}]^{1+}$  signal indicative of two magnetically interacting 4Fe-4S clusters due to features marked with a dotted arrow;  $g$  values marked for rhombic  $[4\text{Fe-4S}]^{1+}$  signal. **d**, X-band continuous wave EPR spectrum of

reduced NpRdhA at 40 K. **e**, X-band continuous wave EPR spectrum of reduced NpRdhA at 80 K showing small base-on cob(II)alamin signal with  $g$  values marked. **f**, Subtraction of **e** from **d** reveals a second axial  $[4\text{Fe-4S}]^{1+}$  signal, with  $g$  values as marked, demonstrating that we can observe EPR signals from two different 4Fe-4S clusters within NpRdhA. Experimental parameters: microwave power 0.5 mW, field modulation frequency 100 KHz, field modulation amplitude 7 G, temperatures as given.



**Extended Data Figure 3 | Crystal structure of NpRdhA.** **a**, Stereo view of an overlay of the NpRdhA cobalamin-binding domain (in blue), the N-terminal non-functional cobalamin-binding domain (in light blue) and the human vitamin B12-processing enzyme CblC (in green). The cobalamin cofactor is bound in a similar base-off manner by both NpRdhA and CblC. The NpRdhA non-functional cobalamin-binding domain does contain an irregular water-filled cavity corresponding to the cofactor binding site, but no longer contains any of the residues implied in cofactor binding. **b**, Representation of the NpRdhA cofactors in space-filling spheres. The second [4Fe-4S] cluster is in van der Waals contact with the corrin ring; the B pyrrol *d*-amide corrin moiety is lined along one side of the cluster and forms a hydrogen bond to one of the S ions. **c**, Detailed view of the NpRdhA fifth ligand-binding site. Residues in

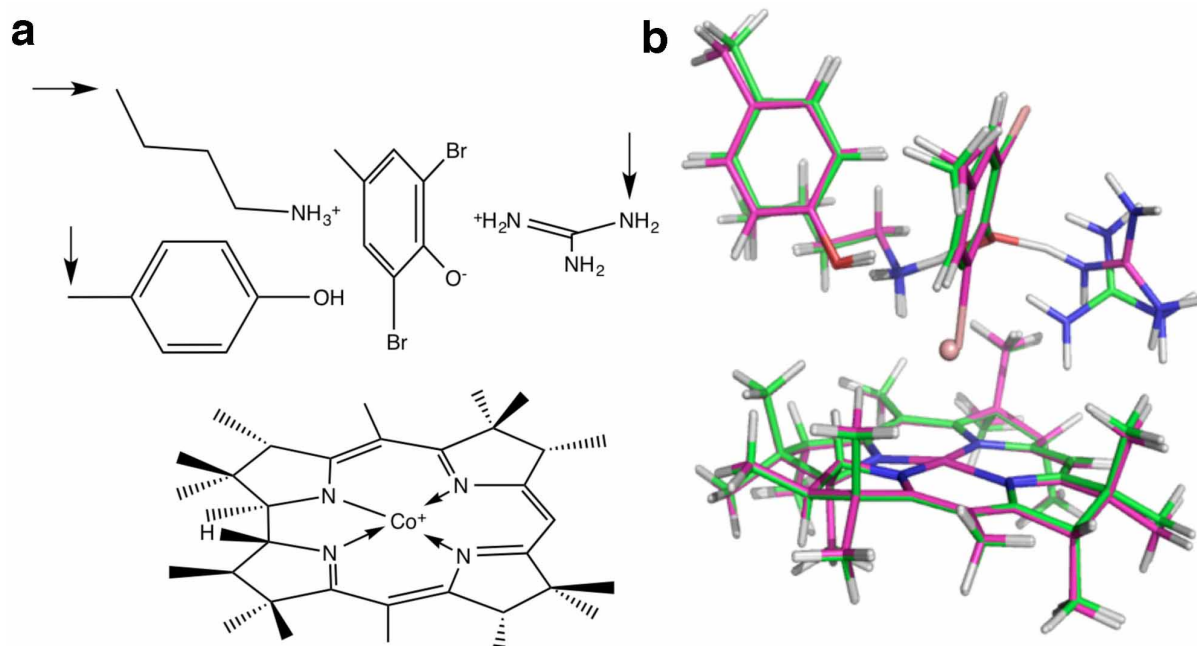
close contact with the chloride ligand are shown in atom coloured sticks (colour coded as in Fig. 2). The  $2F_o - F_c$  map is contoured at  $1\sigma$  and shown as a blue mesh. The Tyr 426-Lys 488 distance is 2.7 Å. **d**, Model of the NpRdhA-substrate complex (colour coded as in Fig. 2e) with the  $F_oF_c$  omit map for an iodide-soaked crystal (to 3.5 Å; contoured at  $6\sigma$ ) depicted by a green mesh. Difference density can be clearly seen at positions corresponding to those predicted to accommodate the bromide atoms of the substrate. **e**, Putative structure of an organometallic substrate-cobalamin intermediate within the NpRdhA active site. The substrate was positioned to minimize close contacts with active site residues. Severe clashes can be observed between the substrate and the corrin ring (highlighted by red lines).



#### Extended Data Figure 4 | Additional EPR spectroscopic data.

**a, b**, Identification of contaminating EPR signals in aerobically purified NpRdhA (150  $\mu$ M). **a**, X-band continuous wave EPR spectrum showing a  $[3\text{Fe-4S}]^{1+}$  cluster signal isolated by subtraction of the spectrum recorded at 20 K from that recorded at 12 K. **b**, X-band continuous wave EPR spectrum showing cob(III)amin- $\text{O}_2^{\bullet-}$  (superoxide) signal isolated by subtracting the spectrum recorded at 20 K from that recorded at 30 K followed by subtraction of a proportion of the  $[3\text{Fe-4S}]^{1+}$  signal. Neither of these signals quantitates to more than 4% of the total EPR signal in the protein as isolated. Experimental parameters: microwave power 0.5 mW, field modulation frequency 100 KHz,

field modulation amplitude 5 G, temperatures as given. **c, d**, Binding of 3,5-dichloro-4-hydroxybenzoate to 150  $\mu$ M NpRdhA leads to a spectrum exhibiting multiple  $A_{\perp}^{\text{Co}}$  and  $A_{\parallel}^{\text{Co}}$  splitting of the spectrum and no apparent chlorine superhyperfine coupling ( $A_{\parallel}^{\text{Cl}}$ ) (spectrum **d**). This suggests relatively disordered binding and possibly substrate movement within the active site on the nanosecond timescale even at the cryogenic temperatures employed in the EPR experiment. Such disorder and dynamics may explain the poor efficacy of this substrate relative to the dibromo analogue (see Fig. 1e) in addition to precluding analysis of the EPR spectrum.

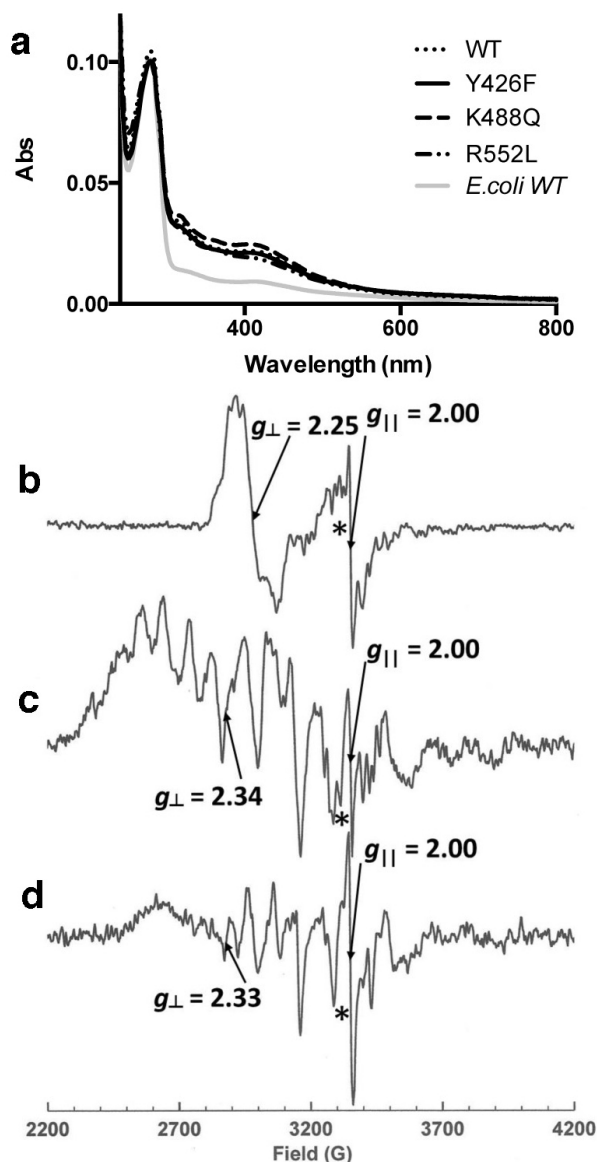


**Extended Data Figure 5 | Structure of the NpRdhA active site DTF model.**

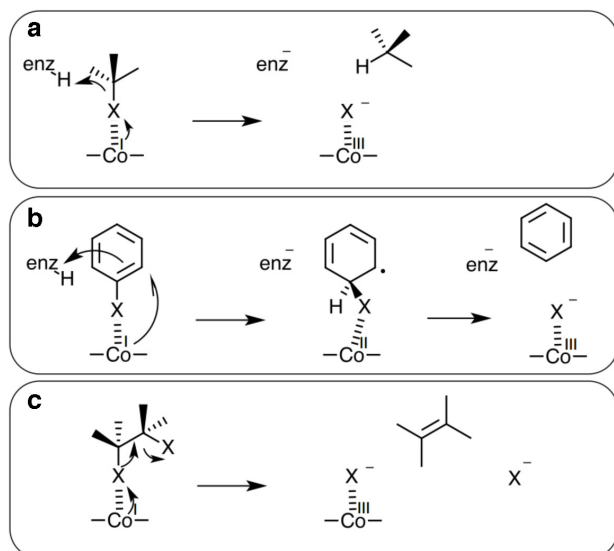
**a, b,** Chemical structures used for DFT calculations (**a**) and overlay of the optimized active site models in the Co(II) (pink carbons) and Co(I) (green carbons; Br is shown as a discrete sphere) oxidation states (**b**). The Co and the three atoms indicated with arrows (**a**) were fixed during optimization. Selected

parameters are given in Extended Data Table 2 and Cartesian coordinates of the optimized structures are included in Supplementary Table 1. The Cbl–Br models comprise the ‘trimmed’ cobalamin with a single axial Br ligand and contain 91 atoms (40 heavy atoms). The full active site model contains 148 atoms (66 heavy atoms).





**Extended Data Figure 6 | Characterization of NpRdhA mutants.** **a**, UV-visible spectra of NpRdhA variants normalized using the  $A_{280}$  absorbance. Mutants Y426F (107  $\mu$ M), K488Q (100  $\mu$ M) and R552L (95  $\mu$ M) have a similar profile as the wild-type enzyme (WT, 176  $\mu$ M) purified from *B. megaterium*. NpRdhA wild type purified from *E. coli* (lacking the corrinoide cofactor; 250  $\mu$ M) is shown for comparison. **b–d**, Continuous wave X-band EPR at 30 K for NpRdhA mutants. **b**, R552L mutant (75  $\mu$ M); **c**, K488Q mutant (65  $\mu$ M); **d**, Y426F mutant (65  $\mu$ M). Each shows the presence of cob(II)alamin which is base on in R552L and base off in the other two mutants. Experimental parameters: microwave power 0.5 mW, modulation frequency 100 KHz, modulation amplitude 7 G.



**Extended Data Figure 7 | Mechanistic proposal for biological reductive dehalogenation.** Our proposed NpRdhA mechanism is fundamentally different from those for other B12-containing enzymes<sup>13,14</sup>, but also differs from the hydrolytic dehalogenases, which use an S<sub>N</sub>2 mechanism whereby either an activated water molecule or a catalytic Asp residue attacks the substrate carbon<sup>5</sup>. The hydrolytic dehalogenases contain a specific halogen-binding site that is believed to contribute to stabilization of the transition state and to facilitate departure of the halide leaving group. In contrast, we propose that the reductive dehalogenase uses the cobalamin cofactor to attack the substrate halogen atom itself, leading to C-halogen bond breakage concomitant with protonation of the leaving group. Distinct variations on this theme could occur within the RdhA family: aliphatic organohalide reductases could operate via heterolytic C-halogen bond cleavage concomitant with halogen-Co(III) bond formation and substrate protonation (a). Those acting on (unactivated) aromatic organohalides are likely to operate via a radical substrate intermediate using homolytic C-halogen bond cleavage (b), while certain reductive dehalogenases have been shown to catalyse vicinal reduction or dihalo-elimination, and we propose that formation of the Co-halogen bond occurs concomitant with leaving of the vicinal halogen atom (c).

**Extended Data Table 1 | Data collection, phasing and refinement statistics**

	RdhA <sub>NP</sub> *
<b>Data collection</b>	
Space group	C2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	177.34, 171.51, 109.02
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 98.87, 90
Resolution (Å)	76.9-2.3 (2.36-2.3)**
<i>R</i> <sub>pim</sub>	6.0 (65.0)
<i>I</i> / $\sigma$ <i>I</i>	9.2 (1.5)
Completeness (%)	98.8 (99.6)
Redundancy	3.3 (3.2)
<b>Refinement</b>	
Resolution (Å)	76.9-2.3 (2.36-2.3)
No. reflections	133721 (9993)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	18.2/20.66 (31.2/35.4)
No. atoms	
Protein	16149
Ion	6
Water	267
B-factors	
Protein	25.4
Ligand/ion	49.1
Water	45.5
R.m.s deviations	
Bond lengths (Å)	0.020
Bond angles (°)	2.07

\* A single crystal was used for data collection.

\*\* Highest resolution shell is shown in parenthesis.

Extended Data Table 2 | EPR parameters for the spectra shown in Fig. 3

Ligand	$g_{\perp}$	Co(II) $A_{\perp}$ (G)	$g_{\parallel}$	Co(II) $A_{\parallel}$ (G)	Ligand $A_{\parallel}$ (G)
H <sub>2</sub> O	2.33	70	2.00	143	-
Cl <sup>-</sup>	2.33	71	2.00	143	25.4
Br <sup>-</sup>	2.30	79*	1.88	139	-
3,5-dibromo-4-hydroxybenzoic acid	2.37	82*	2.00	160	39, 35

\* Estimated only owing to extensive superhyperfine splitting.



Extended Data Table 3 | Selected parameters obtained from DFT calculations of model compounds

	Substrate <sup>a</sup>	Cbl-Br <sup>b</sup>			Active site model <sup>f</sup>	
		Co(I)	Co(II)	Co(III)	Co(I)	Co(II)
energy, Ha	-5488.955	-5500.405	-5500.371	-5500.191	-9184.368	-9184.097
Total atomic charges	Total	-1	-1	0	1	2
	Co		0.56	0.67	0.58	0.72
	Br		-0.76	-0.41	-0.06	-0.05
	Cbl		-0.81	-0.26	-0.82	-0.21
	Substrate <sup>c</sup>				-0.16	-0.02
	Active site <sup>c</sup>				1.47	1.55
Total spin density	Total	0	0	1	0	1
	Co			0.89		1.02
	Br			0.22		0
	Cbl <sup>c</sup>			-0.11		-0.04
	Substrate <sup>c</sup>					0.02
	Active site <sup>c</sup>					0
Bond length <sup>d</sup> , Å	Co-Br		5.19 <sup>e</sup>	2.53	2.72	2.84
	C <sub>S1</sub> -Br	1.94			2.02	1.96
	C <sub>S3</sub> -Br				1.92	1.92
	C <sub>S2</sub> -O <sub>S</sub>				1.34	1.33
	O <sub>Y</sub> -C <sub>S1</sub>				3.11	3.17
	N <sub>K</sub> -O <sub>S</sub>				2.63	2.59
	N <sub>R</sub> -O <sub>S</sub>				2.67	2.59
	Angle, °				174.3	178.2

<sup>a</sup>Deprotonated 2,6-dibromo-4-methylphenol substrate analogue.<sup>b</sup>The trimmed base-off cobalamin with an axial Br ligand (Extended Data Fig 5).<sup>c</sup>The cobalamin moiety excluding the Co atom; the substrate excluding the Br bonded to the Co; the 3 side-chain moieties within the active site model.<sup>d</sup>C<sub>S</sub> and O<sub>S</sub> denote the ring carbons and oxygen, respectively, in the substrate; C<sub>S3</sub>-Br is the bond length of Br not bonded to the Co; O<sub>Y</sub> is the Tyr oxygen; N<sub>K</sub> is the Lys terminal nitrogen; N<sub>R</sub> is the Arg nitrogen closest to the substrate oxygen.<sup>e</sup>Geometry optimization of the Co(I) Cbl-Br species was successful, but the Co-Br bond is broken and the position of the Br is likely to be artefactual.<sup>f</sup>The active site model shown in Extended Data Fig 5. No stable conformation of the Co(III) species was found. If the Tyr 426 proton is transferred to the substrate, then the substrate/product moves away from the cobalamin and no stable structure is observed.

# CAREERS

**TURNING POINT** Chemistry skills apply in fashion design **p.519**

**FAMILY LIFE** Having children while earning a doctorate **go.nature.com/a6ljzt**

**NATUREJOBS** For the latest career listings and advice **www.naturejobs.com**



## TIME MANAGEMENT

# Seize the moment

*A successful leap from postdoc to lab head requires tight control over time and tasks.*

BY JEFFREY M. PERKEL

For the brand-new principal investigator (PI), excitement can soon give way to stress as the mountain of tasks becomes clear. In addition to the main charge — to develop and carry out a research programme, complete with writing grants and managing a lab — the new PI must find time to prepare and

give lectures, grade papers, develop and administer examinations and fulfil service obligations to the institution and scientific community.

It is all a far cry from a postdoctoral fellowship, in which the job is to focus on research and nothing but. The transition can be like jumping from musical soloist to symphony director, says biologist Jing-Ke Weng, who was hired in 2013 as an assistant professor of biology at the

Massachusetts Institute of Technology (MIT) in Cambridge. “You’re not just responsible for your own playing, you are responsible for the music coming from a group of people,” he says.

But new faculty members can adopt strategies to minimize panic and maximize productivity. Some take leadership and management seminars (see ‘Learn the ropes’). Some seek guidance from mentors and administrators, and many rely on classic techniques such as goal-setting and calendar management. PIs need to develop a balancing act that works for them, reserving sufficient time for teaching and service while still focusing on their research. “If things don’t seem to you to be critical for that mission, then deprioritize them,” says Kevan Shokat, chairman of the Department of Cellular and Molecular Pharmacology at the University of California, San Francisco. “Let the department chair deal with the consequences of that.”

Inexperienced PIs often find that it is difficult to establish or to settle into a daily routine, in part because challenges crop up all the time. “Every day has been completely different and, to some extent, unpredictable,” says Marcelo Behar, who last September became a PI in biomedical engineering at the University of Texas at Austin. In a given day Behar might have to negotiate with vendors, meet with students, install laboratory equipment or IT infrastructure, handle administrative or human-resources tasks (such as ensuring that students receive promised tuition waivers) or prepare lectures — and of course, do science.

It helps to maintain a detailed calendar and set aside blocks of time each day and week for tasks such as grant writing, class preparation, keeping up with the literature and lab work. “Manage your time by becoming regimented in your time,” says Joseph Steinmetz, provost at The Ohio State University in Columbus. With experience, that kind of discipline becomes less necessary, he adds.

Behar originally organized his day like that, he says, but the approach proved too inflexible. Dealing with vendors, for instance, could consume 4–6 unanticipated hours per week at the outset. Today, his schedule blends structured and unstructured time, which he manages using Google Calendar. “It took me nearly a semester to figure out a balance that works for me, my students and the department,” he says.

For grant writing, it also helps to break the process down into intermediate deadlines to allow ample time for colleagues to comment on the text and for administrators to approve it, says Julia Salzman, a biochemist at Stanford ▶

ALBIBERON/SHUTTERSTOCK

## HOW TO GET STARTED

*Learn the ropes*

Junior and newly hired faculty members can hone key skills at a wide variety of seminars hosted by universities and funding agencies. Susan Olesik, head of chemistry and biochemistry at The Ohio State University in Columbus, sends her new faculty members to workshops on grant writing and teaching, such as the New Faculty Workshop run by the Cottrell Scholars Collaborative in Washington DC.

The European Molecular Biology Organization (EMBO) in Heidelberg, Germany, offers a course for postdocs that offers tips on managing conflict, communication, building a team and setting goals and priorities, among other issues in laboratory management. Any postdoc with at least 6 months' experience can attend, says programme manager Anne-Marie Glynn. EMBO fellows, young investigators and installation grantees can attend for free; others must pay a fee.

The course advises researchers to divide tasks into four groups according to urgency and importance.

Those that are important and urgent must be tackled immediately. Important but not urgent tasks, such as making progress on a grant application that is still some way from its deadline, should be handled steadily to avoid becoming urgent. Other tasks, such as those delegated by a committee chair, may be considered urgent but not important, and still others as neither important nor urgent.

Ibrahim Cissé, a physicist at the Massachusetts Institute of Technology in Cambridge, took the EMBO course after a three-year postdoc. He says that he found it particularly helpful in helping him to identify the underlying source, motivations and possible solutions for personal conflicts within his team as quickly as possible. "No one budgets time for conflict management," he notes.

EMBO also offers a course for young female PIs that addresses negotiation, leadership and how to say no. Glynn says that these are skills that many must learn to apply, especially when issues of collegiality and institutional hierarchy are involved. At the end of the day, people have to prioritize. "If there was one plan that worked for everyone," she says, "it would be well known." **J.M.P.**



VALERIE NIES

**Cancer researcher Amy Brock (centre) appreciates administrative and technical support in the lab.**

► University in California. She suggests creating faux deadlines by counting backwards from the grant's final deadline. For instance, if the deadline is 1 April, have a draft ready for colleagues' comments by 1 March and a final draft ready for administrators two weeks later.

Those with teaching responsibilities can work on ways to cut preparation time. Susan Olesik, head of chemistry and biochemistry at Ohio State, says that senior faculty members in her department offer lecture notes to younger PIs. Junior PIs can also ask colleagues for guidance on how to gauge the amount and depth of material to cover in the course, how quickly it should be delivered and the degree of difficulty to aim for in preparing exams.

Bonnie Baxter, director of the Great Salt Lake Institute at Westminster College in Salt Lake City, Utah, says that junior faculty members at Westminster are usually required to teach five courses a year, but that two sections of the same class can count as separate courses. Because teaching responsibilities at Westminster are planned up to a year in advance, flexibility may be limited, she says. Still, she recommends discussing it with the department chair or dean. Baxter was also able to count lab-mentoring time — she had accepted five undergraduates into her lab in her first year on the job — as a four-credit research course. "That was one course I didn't have to teach," she says. Further, working with so many eager protégés helped her to advance her research programme.

To bank more precious minutes, Baxter also offers office hours by appointment only, and telecommutes one day a week to minimize interruptions. "Scheduling is everything," she says.

Researchers must also carve out time for other commitments that come with academic positions, including committee work and external roles such as conference organization and advisory-board membership. In many departments, new PIs can decline such requests without penalty, but they cannot say no to

everything. Plus, the more they accept, the more visible they will be, says Hu Li, a computational biologist at the Mayo Clinic in Rochester, Minnesota. It makes sense, he adds, to look at how well those requests dovetail with, or benefit, the research programme. "That's a priority," he says.

Behar, for instance, has just one appointment, to his university's graduate admissions committee. "I get to influence the selection process and make sure candidates who are a good match for my lab get offered admission," he says. Similarly, MIT physicist Ibrahim Cissé has served on a colloquium planning committee and two recruitment panels, assignments that could yield collaborators and students. "There are some committees that are directly related to your research and your research success, and maybe those are the ones that you want to focus on," he says. And always be polite when refusing such requests — cite, for example, a lack of time to provide quality service or a need to learn relevant rules and regulations. A new PI might also offer to defer service for a year rather than refusing a request outright.

**TOKEN MEMBER**

Women sometimes find themselves fielding more requests than their male counterparts, says Anne-Marie Glynn, programme manager for courses and workshops at the European Molecular Biology Organization (EMBO) in Heidelberg, Germany. In departments that have fewer female than male faculty members, women may find that they become the 'token' female member on committees, she says. So it is important to weigh up the benefits of serving.

The department chair (or faculty mentorship committee) can be an ally and a sounding board when deciding which service requests to accept. "I can't stress enough the role of the department chair in this," says Rebecca Ropers-Huilman, chair of the faculty governance executive body at the University of Minnesota in Minneapolis. For instance, some service roles sound easy but



actually require a lot of time and work. Others can yield large benefits for little effort, and a good chairperson can help younger faculty members to figure out which is which.

Shokat, for instance, recommends that junior investigators avoid becoming a PI on large multi-investigator grants because it typically involves a huge time commitment. Instead, he says, it is best to negotiate to be co-PI to tap into the awarded funding and research resources without such a heavy administrative burden.

For those who can swing it, an administrative assistant or grants or lab manager can be tremendously helpful. Amy Brock, a cancer researcher in the biomedical-engineering department at the University of Texas at Austin, says that such staff members can save PIs precious time — for instance, by booking travel, formatting and uploading documents for grant and manuscript submissions, updating CVs and creating and reconciling budgets. “I would advise new faculty to see how administrative support is used by other PIs in the department and then take advantage,” she says.

At most universities, says Shokat, few junior faculty members will get access to administrative assistants. If they do, the person is typically shared. (Shokat’s department gives each faculty member one-third of an assistant’s time.) Glynn recommends that PIs ask whether such resources can be locked in during their hiring negotiations.

Some junior PIs can use start-up funds to pay for an assistant. Kimberly Reynolds, a systems biologist at the University of Texas Southwestern Medical Center in Dallas, hired a lab manager-cum-technician when she started her post. The assistant orders lab equipment and supplies, reconciles budget statements and maintains a lab-reagent archive system. “That’s really taken a lot of stuff off my plate,” Reynolds says.

And, say seasoned PIs, junior faculty members should keep in mind that effective time management includes making time for life outside the lab. Behar promised his wife that he would work 12-hour days four days a week and be home earlier on the fifth. “I’ve been trying to stick with that even though it feels wrong,” he says. Such arrangements often become harder to maintain as tenure decisions approach and stress levels rise. But young investigators also need to remember what they have accomplished and to appreciate where they are. “You’ve been looking to do this for most of your academic career,” says Reynolds. “You finally get to do it, so you should enjoy it.” ■

**Jeffrey M. Perkel** is a freelance writer in Pocatello, Idaho.

## TURNING POINT

# Swati Padmaraj

*Swati Padmaraj was fulfilling her lifelong dream of becoming a chemist when she earned a master’s degree in inorganic chemistry from the University of Bombay, now known as the University of Mumbai, in India. But today she owns and operates a fashion design company, Atiz Fashion House near Seattle, Washington. Here, Padmaraj explains how chemistry continues to play a prominent part in her life.*



### What inspired your interest in chemistry?

I grew up in Mumbai, India, where career choices for those graduating from university were very limited. In the early and mid-1980s, you became a science major, an engineer or a doctor — that was the Indian culture. I was fascinated by chemistry because I could connect minerals and metals with elements in nature. The brilliant chalcopyrite, a copper iron sulfide, is bright yellow, whereas turquoise, a copper aluminium phosphate, is blue-green. The colours that are frequently associated with gems — such as emerald, ruby and sapphire — reflect the mixture of metals in the minerals.

### Did you work as a chemist?

I did a lot of fieldwork in inorganic chemistry while I was pursuing my degrees. Inorganic chemistry is about elements and metals, understanding their molecular structures and composition, and how they affect the environment. I explored the prevalence of metals in different types of rock and soil. There are more than 90 metals in the periodic table and I studied how they coexist in nature.

### Why did you choose to do a master’s degree in business administration instead of a PhD in chemistry?

After my chemistry master’s, I was going to do more research and pursue a PhD. But the chemicals I was using weren’t very safe — I was losing a lot of hair and my skin wasn’t doing well. I decided not to do a PhD, but still wanted a career that involved chemistry. The next best thing I could think of was to get a business degree to find other ways of making that happen.

### What led you to the United States?

My husband was working as an orthodontist in the Seattle area. We decided to raise our family in the United States, and I put my career on hold while I was raising children.

### How did you transition from a degree in inorganic chemistry to fashion design?

I found my passion for creating things when I started thinking about how chemicals affect fabrics. I knew that cotton was the number-one fabric in India, but growing it requires so many pesticides, which pollutes the environment. I also know that polyester can be recycled and can be a good alternative to other fabrics.

### How does your chemistry background influence your fashion design?

Chemistry and fashion are interlinked in many ways. My knowledge of chemistry helps me to choose environmentally friendly fabrics, but it also helps me to understand the environmental implications of various inks that are used for digital printing. Water-soluble inks are good for the environment, but can need more energy to fixate. Also, the chemical make-up of a fabric helps to determine how the fibres will accept colour saturation and at what point the volume of ink might compromise the fabric. My background also helps me to produce garments that last.

### What lies ahead for your design business?

There is so much progress going on in fashion these days, especially when it comes to printing. Block printing requires water washing, which is not as environmentally friendly as digital printing. I’m always looking at how to be environmentally friendly when choosing fabrics. I am using silk and polyester prominently in my current collection, which includes such statement pieces as trench gowns — long-sleeved, hooded dresses — and jumpsuits, both of which are good for people who travel globally. ■

INTERVIEW BY SCOTT KRAFT



# TREATMENT NAIVE

*A testing time.*

BY STEVE ZISSON

The newcomer spotted the cowboy, one arm draped on his horse, bracing himself. He looked as if he would fall over if the horse moved just a few inches.

The newcomer noted the cowboy's weariness and approached with caution.

Not looking up, the cowboy said: "What brings you to town? Got a name?" Before the newcomer could answer, the cowboy sized up his too-new cowboy clothes and said: "I'm going to call you Airman. Simple as that. You popped out of thin air on me."

The cowboy straightened up and seemed to be revived by Airman's approach. A good sign.

Airman shuffled his feet and noticed the brown patina that coated his formerly antiseptic white shoes, the only thing he hadn't replaced when he arrived. No time to get fitted for boots. "I'll take that. Airman. It fits."

Newly named Airman sighed, but knew he had made a connection. The interviews always went much smoother after that. The prize was big and Airman would have to deliver.

"Mind if I asked you some questions?" Airman said as he tapped his tablet. He kept it hidden in his palm, a technique he had learned long ago so as not to spook interviewees.

The cowboy stroked his horse and shrugged. "Day's just about over. Been a hard ride. Wouldn't mind the company. So shoot."

The cowboy leaned against a fence post and chuckled as he rested his hand on his gun.

Airman jumped to attention, a little twitchy about the cowboy's firearm. He reluctantly trained his eyes on the script. "Are you between 18 and 65 years old, inclusive?"

"Yep. On the older end."

"You ever been diagnosed with major depressive disorder, also known as clinical depression or major depression, by a medical doctor?" Got to get better writers for these scripts. He'd tell his boss that later.

The cowboy smiled. "That's a mouthful, Airman. And kind of personal."

Airman would have to go off script. Improvise. "Ever been sad?"

The cowboy scanned the horizon for the sinking sun. "Comes on me almost every day at sunset."

"Good. So it never goes away, this sadness?"

The cowboy tipped his hat and shook his



head. "Get over it when the sun goes down and I hit the sack."

Airman moved on to the next question on his screen. "How tall are you?"

The cowboy measured himself against his horse's shoulder, holding his hand above his head. "Oh, about five foot and a half," he said.

"Your weight?"

"Pretty scrawny, if I do say so. Maybe a buck forty, after Sunday dinner."

The questions would get harder. "Are you taking any antidepressant medication?"

"Shot of whiskey when I need it."

Airman was surprised at the quick response. "Fine," he said. "Alcohol's a depressant. How many drinks do you have a week?"

"Don't usually count them."

Airman sighed. "Three to six? Seven to ten?"

"That's about right."

"I guess that's a seven to ten."

The cowboy looked to the horizon. The sun was pinking out. Airman followed up: "Does whiskey make you feel better?"

"When I'm drinking it, but not the next morning." The cowboy spat a brown gob of tobacco juice, which splashed Airman's shoes.

"Use tobacco?"

The cowboy smiled.

"Doesn't disqualify you," said Airman, who checked a box.

"Do you suffer or have you ever suffered from one of the following?" Airman scanned the list and thought the script writers could have tailored it better to this population.

"Psychotic symptoms? Stroke? Head trauma? Dementia?"

The cowboy considered the list, scratching three days of beard. "Maybe I fell off my horse a few times on my head."

"Doesn't mean you suffered head trauma. Did you go to the doc?"

"Why bother? No blood. Got used to it."

"Okay with me," Airman said, checking off the box.

The big question. "Have you participated in a clinical trial in the past six months?"

The cowboy looked wistful. "Sure have. Took me in an air machine way up there and back east to a big hospital in Boston. Seen your kind before, Airman."

Airman slumped.

"You two were real similar," the cowboy said. "He said it was a hard business, finding patients who were, what did he call it, *treatment naive*. Guess you don't have many left where you come from with all the medications you got."

Airman put away his tablet. "Huh, I figured I was the first in this sector."

The cowboy settled his horse. "I'd do it again for you. You're a nice enough fella. And it was worth getting poked and prodded for the gold pieces and the big ride in the sky."

"Too late. You're not naive anymore. I have to move on."

The messages from his boss had piled up while he was chasing the cowboy. Airman buzzed him.

"You've got to come through for me," the boss said. "This is a blockbuster trial. Time is money. Could be the biggest antidepressant ever."

"You told me this was a pristine time," Airman said. "I'm doing the best I can. CanadaPharm got here first for an anti-anxiety drug trial and got all the eligible volunteers."

"You are giving *me* anxiety," the boss said. "Then give me a better territory," he snapped.

"I've got one," the boss said. "But this is your last chance. Don't blow it."

"Where?" Airman asked.

"I am sending you to China, to a time before they developed medicine. It's an untapped time. Get at least one volunteer or you're through." ■

Steve Zisson is a journalist who has written extensively about clinical research.

ILLUSTRATION BY JACEY

➔ NATURE.COM

Follow Futures:

@NatureFutures

go.nature.com/mtoodm